



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **2000285253 A**(43) Date of publication of application: **13.10.00**

(51) Int. Cl.

G06T 13/00
G06F 17/30
// H04N 5/275
H04N 7/32

(21) Application number: **11187033**(22) Date of filing: **30.06.99**(30) Priority: **28.01.99 JP 11020387**(71) Applicant: **TOSHIBA CORP**

(72) Inventor: **KANEKO TOSHIMITSU**
HORI OSAMU
MITA TAKESHI
YAMAMOTO KOJI

(54) **METHOD FOR DESCRIPTION OF OBJECT AREA INFORMATION, GENERATOR OF OBJECT AREA INFORMATION, METHOD FOR VIDEO INFORMATION PROCESSING, AND INFORMATION PROCESSOR**

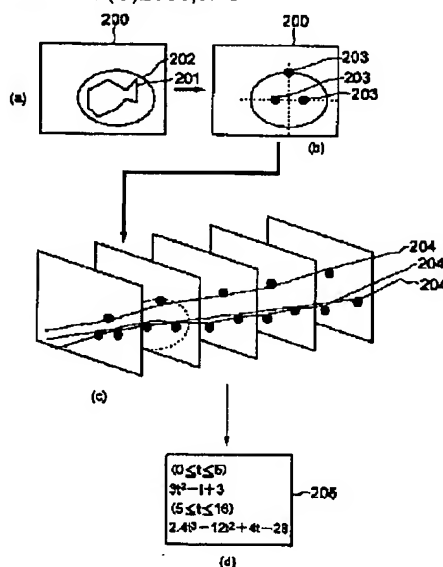
prescribed function 204, and the parameter 205 of this function is used to describe information related to the area of this object.

COPYRIGHT: (C)2000,JPO

(57) Abstract:

PROBLEM TO BE SOLVED: To realize an object area description method by which the area of a desired object in a picture is described in a smaller data volume and its generation and data handling are facilitated.

SOLUTION: In this object area information description method where information related to the area of an arbitrary object in the picture is described over a plurality of consecutive frames, an area 201 of the object in the picture is specified by at least one of an approximate figure 202 of the area of this object and characteristic points, and a locus obtained by arranging the positional data of representative points 203 of the approximate figure 202 or of characteristic points in accordance with advance of frames 200 is approximated by a



(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2000-285253

(P2000-285253A)

(43)公開日 平成12年10月13日(2000. 10. 13)

(51)Int.Cl. ⁷	識別記号	F I	テ-マコ-ト*(参考)
G 0 6 T 13/00		G 0 6 F 15/62	3 4 0 A 5 B 0 5 0
G 0 6 F 17/30		H 0 4 N 5/275	5 B 0 7 5
// H 0 4 N 5/275		G 0 6 F 15/419	3 2 0 5 C 0 2 3
7/32		15/40	3 7 0 G 5 C 0 5 9
		H 0 4 N 7/137	Z
		審査請求 未請求 請求項の数25 O L (全 31 頁)	

(21)出願番号 特願平11-187033

(22)出願日 平成11年6月30日(1999. 6. 30)

(31)優先権主張番号 特願平11-20387

(32)優先日 平成11年1月28日(1999. 1. 28)

(33)優先権主張国 日本 (J P)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 金子 敏充

神奈川県川崎市幸区小向東芝町1番地 株

式会社東芝研究開発センター内

(72)発明者 堀 修

神奈川県川崎市幸区小向東芝町1番地 株

式会社東芝研究開発センター内

(74)代理人 100058479

弁理士 鈴江 武彦 (外6名)

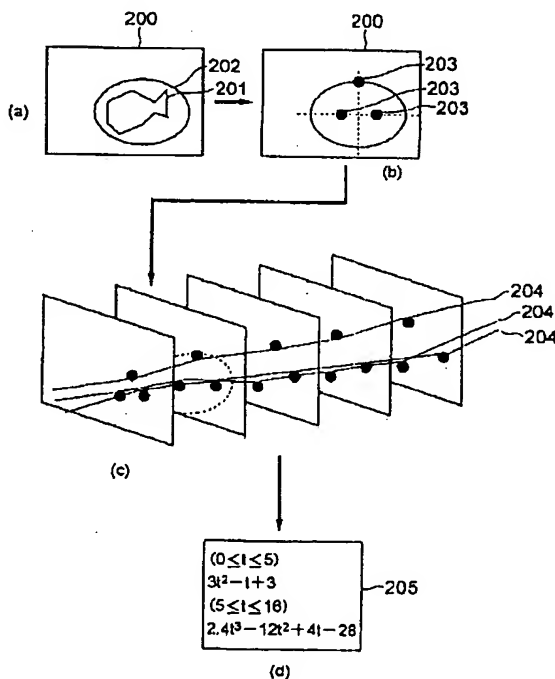
最終頁に続く

(54)【発明の名称】 物体領域情報記述方法及び物体領域情報生成装置並びに映像情報処理方法及び情報処理装置

(57)【要約】

【課題】 映像中の所望の物体の領域を少ないデータ量で記述でき且つその作成やデータの扱いも容易にする物体領域記述方法を提供すること。

【解決手段】 映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、映像中における対象となる物体の領域201を、該物体の領域に対する近似図形202もしくは特徴点の少なくとも一方で特定し、該近似図形202の代表点203もしくは該特徴点の位置データをフレーム200の進行に沿って並べたときの軌跡を所定の関数204で近似し、該関数のパラメータ205を用いて該物体の領域に関する情報を記述することを特徴とする。



1

【特許請求の範囲】

【請求項 1】映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、

映像中における対象となる物体の領域を、該物体の領域に対する近似図形もしくは特徴点の少なくとも一方で特定し、該近似図形の代表点もしくは該特徴点の位置データをフレームの進行に沿って並べたときの軌跡を所定の関数で近似し、該関数のパラメータを用いて該物体の領域に関する情報を記述することを特徴とする物体領域情報記述方法。

【請求項 2】映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、
少なくとも、対象となる前記複数フレームのうちの先頭のフレームの番号及び最終のフレームの番号もしくは先頭のフレームのタイムスタンプ及び最終のフレームのタイムスタンプを特定可能な情報と、該物体の領域を近似する近似図形の図形種を識別する情報と、該物体の領域に対する近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとを用いて、前記物体の領域に関する情報を記述することを特徴とする物体領域情報記述方法。

【請求項 3】映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、
少なくとも、対象となる前記複数フレームのうちの先頭のフレームの番号及び最終のフレームの番号もしくは先頭のフレームのタイムスタンプ及び最終のフレームのタイムスタンプを特定可能な情報と、該物体の領域を近似する近似図形の数と、該近似図形の図形種を識別する情報と、各近似図形に対応する、該近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとを用いて、前記物体の領域に関する情報を記述することを特徴とする物体領域情報記述方法。

【請求項 4】映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、
少なくとも、対象となる前記複数フレームのうちの先頭のフレームの番号及び最終のフレームの番号もしくは先頭のフレームのタイムスタンプ及び最終のフレームのタイムスタンプを特定可能な情報と、該物体の領域の特徴点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとを用いて、前記物体の領域に関する情報を記述することを特徴とする物体領域情報記述方法。

【請求項 5】前記関数のパラメータは、前記軌跡の節点の位置データおよび該節点の位置データとともに用いて前記軌跡を特定可能な情報であることを特徴とする請求

2

項 1 ないし 4 に記載の物体領域情報記述方法。

【請求項 6】前記物体の領域に対する近似図形の代表点もしくは前記物体の領域の特徴点が複数存在する場合に、現在のフレームにおける複数の代表点もしくは特徴点の各々が互いに隣接するフレームにおける複数の代表点もしくは特徴点のいずれに対応するかを特定することを特徴とする請求項 1 ないし 4 のいずれか 1 項に記載の物体領域情報記述方法。

【請求項 7】前記物体に関連付けられている関連情報または該関連情報へのアクセス方法を示す情報を併せて記述することを特徴とする請求項 1 ないし 4 のいずれか 1 項に記載の物体領域情報記述方法。

【請求項 8】映像データをもとにして請求項 1 ないし 4 のいずれか 1 項に記載の物体領域情報記述方法により記述された 1 または複数の物体の領域に関する情報を含む物体領域データを格納した記録媒体。

【請求項 9】映像データをもとにして請求項 1 ないし 4 のいずれか 1 項に記載の物体領域情報記述方法により記述された 1 または複数の物体の領域に関する情報および各物体に関連付けられている関連情報または該関連情報へのアクセス方法を示す情報を含む物体領域データを格納した記録媒体。

【請求項 10】映像データをもとにして請求項 1 ないし 4 のいずれか 1 項に記載の物体領域情報記述方法により記述された 1 または複数の物体の領域に関する情報および各物体に関連付けられている関連情報を特定するための情報を含む物体領域データと、各物体に関連付けられている関連情報とを格納した記録媒体。

【請求項 11】映像を表示している画面において特定の物体が指定されたか否かを判定するための映像情報処理方法であって、

映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域に対する近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、前記取得した情報に基づいて、該フレームにおける前記代表点の位置を求め、求められた前記代表点の位置をもとにして前記近似図形の位置を求め、前記指定された位置が求められた前記近似図形の内部に存在するか否かを調べ、内部に存在すると判断された場合に前記物体が指定されたと判定することを特徴とする映像情報処理方法。

【請求項 12】映像を表示している画面において特定の物体が指定されたか否かを判定するための映像情報処理方法であって、

映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域の特徴点の位置データをフレーム

3

の進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、

前記取得した情報に基づいて、該フレームにおける前記特徴点の位置を求め、

前記指定された位置と求められた前記特徴点の位置との間の距離が基準値以下であるか否かを調べ、基準値以下であると判断された場合に前記物体が指定されたと判定することを特徴とする映像情報処理方法。

【請求項 13】前記物体が指定されたと判定された場合には、前記物体に関連付けられている関連情報を表示することを特徴とする請求項 11 または 12 に記載の映像情報処理方法。

【請求項 14】映像を表示している画面において特定の物体が指定された場合に該物体の存在する領域を明示的に表示するための映像情報処理方法であって、

前記映像中に予め設定された物体の領域が存在する場合に、該物体の領域に対する近似図形の代表点もしくは該物体の領域の特徴点の少なくとも一方の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、

前記取得した情報に基づいて、該フレームにおける前記代表点または特徴点を求め、

求められた前記代表点または特徴点に基づいて、前記物体の領域の画面上における存在箇所を明示するための情報を所定の表示形態で表示することを特徴とする映像情報処理方法。

【請求項 15】映像中に出現する物体のうち条件に該当するものを検索する映像情報処理方法であって、検索対象となる映像中の任意の位置および該位置に基づく検索条件を入力し、

前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、

前記取得した情報に基づいて求めた 1 つの物体の 1 つのフレームにおける前記近似図形もしくは前記特徴点と、前記指定された位置とが、所定の関係にあるか否かを判定することを、複数のフレームに渡って各物体について行い、

前記判定の結果に基づいて、前記検索条件に該当する物体を求めることを特徴とする映像情報処理方法。

【請求項 16】前記所定の関係は、前記指定された位置が前記近似図形の内部に存在する関係もしくは前記特徴点と前記指定された位置との間の距離が基準値以下である関係であり、

前記検索条件は、抽出すべき物体の条件として、前記指定された位置に関して前記所定の関係の成立するフレームが少なくとも 1 つ存在する条件、前記指定された位置に関して前記所定の関係の成立するフレームが一定数連

4

続して存在する条件、および前記所定の関係が全てのフレームにおいて成立しない条件を含む検索条件群のなかから選択されたものであることを特徴とする請求項 15 に記載の映像情報処理方法。

【請求項 17】映像中に出現する物体のうち条件に該当するものを検索する映像情報処理方法であって、検索対象となる映像中の位置の軌跡を指定する情報を入力し、

前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、

前記取得した情報に基づいて求めた 1 つの物体の前記代表点もしくは前記特徴点の軌跡と、前記指定された位置の軌跡との類似性を評価することを、各物体について行い、

評価された前記類似性に基づいて、指定された前記軌跡に該当する物体を求めることを特徴とする映像情報処理方法。

【請求項 18】前記位置の軌跡を指定する情報は、位置と時間との関係を含む時系列情報であり、

前記類似性は、時間的な位置関係を加味して評価することを特徴とする請求項 17 に記載の映像情報処理方法。

【請求項 19】映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述したデータを生成するための物体領域情報生成装置であって、対象となる前記複数のフレームについて、映像中における対象となる物体の領域を、所定の図形を用いて近似する手段と、

前記近似に用いられた図形を特定可能な所定数の代表点の複数のフレーム中における座標値を求める手段と、

連続する前記複数フレームに渡って求められた各代表点の各座標値について、該代表点の座標値の時系列の軌跡を所定の関数で近似する手段とを備え、

前記関数のパラメータを用いて前記物体の領域に関する情報を生成することを特徴とする物体領域情報生成装置。

【請求項 20】映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述したデータを生成するための物体領域情報生成装置であって、対象となる前記複数のフレームについて、映像中における対象となる物体の領域に対する所定数の特徴点の座標値を求める手段と、

連続する前記複数フレームに渡って求められた各特徴点の各座標値について、該特徴点の座標値の時系列の軌跡を所定の関数で近似する手段とを備え、

前記関数のパラメータを用いて前記物体の領域に関する情報を生成することを特徴とする物体領域情報生成装置。

5

【請求項 2 1】映像を表示している画面において特定の物体が指定された場合に所定の処理を行う情報処理装置であって、

映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域に対する近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを取得し、該フレームにおける前記代表点の位置を求める手段と、

求められた前記代表点の位置をもとにして前記近似図形の位置を求める手段と、

前記指定された位置が求められた前記近似図形の内部に存在するか否かを判定する手段とを備えたことを特徴とする情報処理装置。

【請求項 2 2】映像を表示している画面において特定の物体が指定された場合に所定の処理を行う情報処理装置であって、

映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域の特徴点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを取得し、該フレームにおける特徴点の位置を求める手段と、

前記指定された位置と求められた前記特徴点の位置との間の距離が基準値以下であるか否かを判定する手段とを備えたことを特徴とする情報処理装置。

【請求項 2 3】映像を表示している画面において特定の物体が指定された場合に所定の処理を行う情報処理装置であって、

前記映像中に予め設定された物体の領域が存在する場合に、該物体の領域に対する近似図形の代表点もしくは該物体の領域の特徴点の少なくとも一方の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを取得し、該フレームにおける代表点または特徴点を求める手段と、

求められた前記代表点または特徴点に基づいて、前記物体の領域の画面上における存在箇所を明示するための情報を所定の表示形態で表示する手段とを備えたことを特徴とする情報処理装置。

【請求項 2 4】映像中に出現する物体のうち指定された条件に該当するものを検索する情報処理装置であって、検索対象となる映像中の任意の位置および該位置に基づく検索条件が入力された際、前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得する手段と、

前記取得した情報に基づいて求めた 1 つの物体の 1 つのフレームにおける前記近似図形もしくは前記特徴点と、

6

前記指定された位置とが、所定の関係にあるか否かを判定することを、複数のフレームに渡って各物体について行う手段と、

前記判定の結果に基づいて、前記検索条件に該当する物体を求める手段とを備えたことを特徴とする情報処理装置。

【請求項 2 5】映像中に出現する物体のうち指定された条件に該当するものを検索する情報処理装置であって、検索対象となる映像中の位置の軌跡を指定する情報が入力された際、前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得する手段と、

前記取得した情報に基づいて求めた 1 つの物体の前記代表点もしくは前記特徴点の軌跡と、前記指定された位置の軌跡との類似性を評価することを、各物体について行う手段と、

評価された前記類似性に基づいて、指定された前記軌跡に該当する物体を求める手段とを備えたことを特徴とする映像情報処理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、映像中の物体の領域に関する情報を記述するための物体領域情報記述方法、映像中の物体の領域に関する情報を生成するための物体領域情報生成装置、並びに映像中の物体に対する指定を受け所定の処理を行うあるいは映像中の物体の検索を行う情報処理装置及びそのための映像情報処理方法に関する。

【0002】

【従来の技術】ハイパーメディアは、映像、音声、テキストなどのメディアの間にハイパーリンクと呼ばれる関連情報を付与し、相互に参照できるようにしたものである。映像を中心にした場合、例えば映像中の登場物体に関連情報が付与されており、この物体が指示されると関連情報の表示を行うというのがハイパーメディアの代表例である。このとき、映像中の物体は映像のフレーム番号もしくはタイムスタンプと映像中の領域を特定する情報とで表現され、映像データの中にもしくは別データとして記録されている。

【0003】映像中の領域を特定する方法としては、マスク画像がよく利用されてきた。これは指定領域内の場合と指定領域外の場合で異なる画素値を与えて構成する画像である。例えば、領域内の場合には 1、領域外の場合には 0 という画素値を与えるのが最も簡単な方法である。また、CG などに使われる α 値を利用することもある。通常、 α 値は 256 階調の値を表現できるので、そのうちの一部を使い、例えば指定領域内の場合には 255、指定領域外の場合には 0 と表現する。このようなマスク画像

7

により画像中の領域が表現されている場合、あるフレームにおける画素が、指定領域内であるかどうかを判定するには、そのフレームに該当するマスク画像の該当画素の値を読み取り、0であるか255であるかにより簡単に判定することができる。マスク画像はどのような形の領域でも、また不連続な領域でも表現できるという自由度を持っているが、画像サイズと同じサイズの画素を持つ必要がある。

【0004】マスク画像のデータ量を削減するために、マスク画像の圧縮がよく利用される。0、1の2値のマスク画像の場合には、2値画像としての処理ができるため、ファクシミリ等で用いられている圧縮方法が利用されることが多い。また、ISO/IEC動画圧縮標準化グループMPEG (Moving Picture Experts Group) が標準化しているMPEG-4では、0、1の2値のマスク画像の他、 α 値を利用したマスク画像までを圧縮対象とした任意形状符号化を採用することになっている。これは、動き補償を用いた圧縮手法であり、圧縮効率が向上するが、その分、圧縮・復号過程は複雑になる。

【0005】以上のように映像中のある領域を表現するにはマスク画像かあるいはマスク画像を圧縮したものを利用することが多かったが、領域を特定するためのデータとしては、より簡単にかつ高速に取り出すことが可能でデータ量も小さく、そして容易に扱うことができる形態のものが望まれている。

【0006】一方、映像中の動く物体の関連情報を表示させるような操作が通常想定されるハイパーメディアにおいては、静止画像を対象とするのとは異なって物体の指示にある程度の困難さを伴うものであり、ユーザが物体の特定の箇所を狙ってそこを指示することは通常難しく、またユーザも通常は物体の例えば中心付近を漠然と狙って指示しようとする場合が多いと考えられ、さらに物体の動きによっては厳密には物体の外部であるような物体の近傍を指示してしまうことも多い。したがって、領域を特定するためのデータとして、このようなメディアにより適合した形態のものが望まれている。また、映像中の動く物体の関連情報を表示するようなシステムにおいては、ユーザによる映像中の動く物体の指示をより容易にするように支援する機構が望まれている。

【0007】

【発明が解決しようとする課題】以上説明したように、映像中の所望の物体の領域を表現する方法として、マスク画像ではデータ量が多くなるという問題があり、また、マスク画像を圧縮して用いると符号化・復号化が複雑になり、しかも直接データを編集することができないためハンドリングが難しいという問題があった。

【0008】また、ユーザによる映像中の動く物体の指示をより容易にするような仕組みが提供されていないという問題があった。

8

【0009】本発明は、上記事情を考慮してなされたもので、映像中の所望の物体の領域を少ないデータ量で記述でき且つその作成やそのデータの扱いも容易にする物体領域情報記述方法及び物体領域情報生成装置を提供することを目的とする。

【0010】また、本発明は、ユーザによる映像中の物体の指示やその判定を容易にする物体領域情報記述方法及び物体領域情報生成装置並びに映像情報処理方法及び情報処理装置を提供することを目的とする。

【0011】また、本発明は、映像中の物体の検索を容易にする物体領域情報記述方法及び物体領域情報生成装置並びに映像情報処理方法及び情報処理装置を提供することを目的とする。

【0012】

【課題を解決するための手段】本発明は、映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、映像中における対象となる物体の領域を、該物体の領域に対する近似図形もしくは特徴点の少なくとも一方で特定し、該近似図形の代表点もしくは該特徴点の位置データをフレームの進行に沿って並べたときの軌跡を所定の関数で近似し、該関数のパラメータを用いて該物体の領域に関する情報を記述することを特徴とする。

【0013】また、本発明は、映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、少なくとも、対象となる前記複数フレームのうちの先頭のフレームの番号及び最終のフレームの番号もしくは先頭のフレームのタイムスタンプ及び最終のフレームのタイムスタンプを特定可能な情報と、該物体の領域を近似する近似図形の図形種を識別する情報と、該物体の領域に対する近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとを用いて、前記物体の領域に関する情報を記述することを特徴とする。

【0014】また、本発明は、映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、少なくとも、対象となる前記複数フレームのうちの先頭のフレームの番号及び最終のフレームの番号もしくは先頭のフレームのタイムスタンプ及び最終のフレームのタイムスタンプを特定可能な情報と、該物体の領域を近似する近似図形の数と、該近似図形の図形種を識別する情報と、各近似図形に対応する、該近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとを用いて、前記物体の領域に関する情報を記述することを特徴とする。

【0015】また、本発明は、映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述するための物体領域情報記述方法であって、少なく

9

とも、対象となる前記複数フレームのうちの先頭のフレームの番号及び最終のフレームの番号もしくは先頭のフレームのタイムスタンプ及び最終のフレームのタイムスタンプを特定可能な情報と、該物体の領域の特徴点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとを用いて、前記物体の領域に関する情報を記述することを特徴とする。

【0016】対象となる前記複数フレームのうちの先頭のフレームの番号及び最終のフレームの番号もしくは先頭のフレームのタイムスタンプ及び最終のフレームのタイムスタンプを特定可能な情報は、例えば、先頭のフレームの番号及び最終のフレームの番号、あるいは先頭のフレームの番号及び先頭のフレームの番号と最終のフレームの番号との差分である。タイムスタンプを用いる場合も同様である。

【0017】好ましくは、前記関数のパラメータは、前記軌跡の節点の位置データおよび該節点の位置データとともに用いて前記軌跡を特定可能な情報であるようにしてもよい。あるいは、前記関数のパラメータは、関数の係数であるようにしてもよい。

【0018】好ましくは、前記物体の領域に対する近似図形の代表点もしくは前記物体の領域の特徴点が複数存在する場合に、現在のフレームにおける複数の代表点もしくは特徴点の各々が互いに隣接するフレームにおける複数の代表点もしくは特徴点のいずれに対応するかを特定するようにしてもよい。

【0019】好ましくは、前記物体に関連付けられている関連情報または該関連情報へのアクセス方法を示す情報を併せて記述するようにしてもよい。

【0020】また、本発明は、映像データをもとにして請求項1ないし4のいずれか1項に記載の物体領域情報記述方法により記述された1または複数の物体の領域に関する情報を含む物体領域データを格納した記録媒体である。また、本発明は、映像データをもとにして請求項1ないし4のいずれか1項に記載の物体領域情報記述方法により記述された1または複数の物体の領域に関する情報および各物体に関連付けられている関連情報または該関連情報へのアクセス方法を示す情報を含む物体領域データを格納した記録媒体である。また、本発明は、映像データをもとにして請求項1ないし4のいずれか1項に記載の物体領域情報記述方法により記述された1または複数の物体の領域に関する情報および各物体に関連付けられている関連情報を特定するための情報を含む物体領域データと、各物体に関連付けられている関連情報とを格納した記録媒体である。

【0021】また、本発明は、映像を表示している画面において特定の物体が指定されたか否かを判定するための映像情報処理方法であって、映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域に

10

対する近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、前記取得した情報に基づいて、該フレームにおける前記代表点の位置を求め、求められた前記代表点の位置をもとにして前記近似図形の位置を求め、前記指定された位置が求められた前記近似図形の内部に存在するか否かを調べ、内部に存在すると判断された場合に前記物体が指定されたと判定することを特徴とする。

【0022】また、本発明は、映像を表示している画面において特定の物体が指定されたか否かを判定するための映像情報処理方法であって、映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域の特徴点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、前記取得した情報に基づいて、該フレームにおける前記特徴点の位置を求め、前記指定された位置と求められた前記特徴点の位置との間の距離が基準値以下であるか否かを調べ、基準値以下であると判断された場合に前記物体が指定されたと判定することを特徴とする。好ましくは、前記物体が指定されたと判定された場合には、前記物体に関連付けられている関連情報を呈示するようにしてもよい。

【0023】また、本発明は、映像を表示している画面において特定の物体が指定された場合に該物体の存在する領域を明示的に表示するための映像情報処理方法であって、前記映像中に予め設定された物体の領域が存在する場合に、該物体の領域に対する近似図形の代表点もしくは該物体の領域の特徴点の少なくとも一方の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、前記取得した情報に基づいて、該フレームにおける前記代表点または特徴点を求め、求められた前記代表点または特徴点に基づいて、前記物体の領域の画面上における存在箇所を明示するための情報を所定の表示形態で表示することを特徴とする。

【0024】また、本発明は、映像中に出現する物体のうち条件に該当するものを検索する映像情報処理方法であって、検索対象となる映像中の任意の位置および該位置に基づく検索条件を入力し、前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、前記取得した情報に基づいて求めた1つの物体の1つのフレームにおける前記近似図形もしくは前記特徴点と、前記指定された位置とが、所定の関係にあるか否かを判定することを、複数のフレームに渡って各物体について行い、前記判定の結果に基づいて、前記検索条件に該当す

11

る物体を求めることを特徴とする。好ましくは、前記所定の関係は、前記指定された位置が前記近似図形の内部に存在する関係もしくは前記特徴点と前記指定された位置との間の距離が基準値以下である関係であり、前記検索条件は、抽出すべき物体の条件として、前記指定された位置に関して前記所定の関係の成立するフレームが少なくとも1つ存在する条件、前記指定された位置に関して前記所定の関係の成立するフレームが一定数連続して存在する条件、および前記所定の関係が全てのフレームにおいて成立しない条件を含む検索条件群のなかから選択されたものであるようにしてもよい。また、好ましくは、前記検索条件群は、さらに、前記位置に基づく条件に加重される条件として、前記物体に対する近似図形が満たすべき属性条件を含むようにしてもよい。

【0025】また、本発明は、映像中に出現する物体のうち条件に該当するものを検索する映像情報処理方法であって、検索対象となる映像中の位置の軌跡を指定する情報を入力し、前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得し、前記取得した情報に基づいて求めた1つの物体の前記代表点もしくは前記特徴点の軌跡と、前記指定された位置の軌跡との類似性を評価することを、各物体について行い、評価された前記類似性に基づいて、指定された前記軌跡に該当する物体を求めることを特徴とする。好ましくは、前記位置の軌跡を指定する情報は、位置と時間との関係を含む時系列情報であり、前記類似性は、時間的な位置関係を加味して評価するようにしてもよい。なお、前記軌跡については、例えば、映像中の物体を指定し、該物体の持つ軌跡を指定された軌跡としてもいいし、あるいは、ユーザが前記軌跡をGUI上で描くように入力可能としてもよい。

【0026】また、本発明は、映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述したデータを生成するための物体領域情報生成装置であって、対象となる前記複数のフレームについて、映像中における対象となる物体の領域を、所定の図形を用いて近似する手段と、前記近似に用いられた図形を特定可能な所定数の代表点の複数のフレーム中における座標値を求める手段と、連続する前記複数フレームに渡って求められた各代表点の各座標値について、該代表点の座標値の時系列の軌跡を所定の関数で近似する手段とを備え、前記関数のパラメータを用いて前記物体の領域に関する情報を生成することを特徴とする。

【0027】また、本発明は、映像中における任意の物体の領域に関する情報を連続する複数フレームに渡って記述したデータを生成するための物体領域情報生成装置であって、対象となる前記複数のフレームについて、映

12

像中における対象となる物体の領域に対する所定数の特徴点の座標値を求める手段と、連続する前記複数フレームに渡って求められた各特徴点の各座標値について、該特徴点の座標値の時系列の軌跡を所定の関数で近似する手段とを備え、前記関数のパラメータを用いて前記物体の領域に関する情報を生成することを特徴とする。

【0028】また、本発明は、映像を表示している画面において特定の物体が指定された場合に所定の処理を行う情報処理装置であって、映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域に対する近似図形の代表点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを取得し、該フレームにおける前記代表点の位置を求める手段と、求められた前記代表点の位置をもとにして前記近似図形の位置を求める手段と、前記指定された位置が求められた前記近似図形の内部に存在するか否かを判定する手段とを備えたことを特徴とする。

【0029】また、本発明は、映像を表示している画面において特定の物体が指定された場合に所定の処理を行う情報処理装置であって、映像を表示している画面において任意の位置が指定された際、該映像中に予め設定された物体の領域が存在する場合に、該物体の領域の特徴点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを取得し、該フレームにおける特徴点の位置を求める手段と、前記指定された位置と求められた前記特徴点の位置との間の距離が基準値以下であるか否かを判定する手段とを備えたことを特徴とする。

【0030】また、本発明は、映像を表示している画面において特定の物体が指定された場合に所定の処理を行う情報処理装置であって、前記映像中に予め設定された物体の領域が存在する場合に、該物体の領域に対する近似図形の代表点もしくは該物体の領域の特徴点の少なくとも一方の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを取得し、該フレームにおける代表点または特徴点を求める手段と、求められた前記代表点または特徴点に基づいて、前記物体の領域の画面上における存在箇所を明示するための情報を所定の表示形態で表示する手段とを備えたことを特徴とする。

【0031】また、本発明は、映像中に出現する物体のうち指定された条件に該当するものを検索する情報処理装置であって、検索対象となる映像中の任意の位置および該位置に基づく検索条件が入力された際、前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得する手段と、前記取得した情報に基づいて求めた

1つの物体の1つのフレームにおける前記近似図形もしくは前記特徴点と、前記指定された位置とが、所定の関係にあるか否かを判定することを、複数のフレームに渡って各物体について行う手段と、前記判定の結果に基づいて、前記検索条件に該当する物体を求める手段とを備えたことを特徴とする。

【0032】また、本発明は、映像中に出現する物体のうち指定された条件に該当するものを検索する情報処理装置であって、検索対象となる映像中の位置の軌跡を指定する情報が入力された際、前記検索対象となる映像中に出現する物体毎に作成された、該物体の領域に対する近似図形の代表点の位置データもしくは該物体の領域の特徴点をフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータを記述した情報を取得する手段と、前記取得した情報に基づいて求めた1つの物体の前記代表点もしくは前記特徴点の軌跡と、前記指定された位置の軌跡との類似性を評価することを、各物体について行う手段と、評価された前記類似性に基づいて、指定された前記軌跡に該当する物体を求める手段とを備えたことを特徴とする。

【0033】なお、装置に係る本発明は方法に係る発明としても成立し、方法に係る本発明は装置に係る発明としても成立する。

【0034】また、装置または方法に係る本発明は、コンピュータに当該発明に相当する手順を実行させるための（あるいはコンピュータを当該発明に相当する手段として機能させるための、あるいはコンピュータに当該発明に相当する機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても成立する。

【0035】本発明によれば、複数フレームに渡る映像中の物体の領域を、該物体の領域に対する近似図形の代表点または該物体の領域の特徴点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとして記述することにより、複数フレームに渡る映像中の物体の領域を少量の関数パラメータのみによって記述することができるため、物体の領域を特定するためのデータの量を効果的に削減することができ、またハンドリングを容易にすることができる。また、近似図形からの代表点あるいは特徴点の抽出や、近似曲線のパラメータの生成も容易に行うことができる。また、近似曲線のパラメータから近似図形を生成することも容易に行うことができる。

【0036】また、近似図形の代表点を用いる場合、この近似図形として基本的な図形、例えば一つまたは複数の楕円を用い、例えば楕円を二つの焦点と他の1点により代表させれば、ユーザにより指定された任意の座標が物体の領域（近似図形）内か否かを簡単な判定式により判定することができる。さらに、これによって、ユーザによる映像中の動く物体の指示をより容易にすることが

できる。

【0037】また、特徴点を用いる場合、ユーザにより指定された任意の座標が物体の領域を指示しているか否かを極めて簡単に判定することができる。さらに、これによって、ユーザによる映像中の動く物体の指示をより容易にすることができる。

【0038】さらに、本発明による物体領域データを用いて特定される映像中の物体の領域のうち関連情報を持つ物体の領域の表示や物体の領域を示す画像の表示を制御することにより、例えばユーザが関連情報があるかなにか、そしてその物体の領域の位置を一目で認識できるようにするなど、ユーザの操作を効果的に支援することができる。

【0039】また、本発明によれば、映像中の物体の通過位置、ある地点での滞留時間、あるいは軌跡などに基づいて、映像中の物体の検索を容易に行うことができる。

【0040】

【発明の実施の形態】以下、図面を参照しながら発明の実施の形態を説明する。

【0041】（第1の実施形態）図1に、本発明の第1の実施形態に係る物体領域データ変換装置の構成例を示す。図1に示されるように、本物体領域データ変換装置は、映像データ記憶部100、領域抽出部101、領域の図形近似部102、図形代表点抽出部103、代表点の曲線近似部104、関連情報記憶部105、領域データ記憶部106を備えている。なお、本処理（特に、領域抽出部101や領域の図形近似部102の処理）においてユーザの操作を介入させる形態を取る場合には、映像データを例えばフレーム単位で表示させ、ユーザの指示入力等を受け付けるGUIが用いられる（図1では省略している）。

【0042】映像データ記憶部100は、映像データが記憶されているもので、例えばハードディスクや光ディスク、半導体メモリなどで構成される。

【0043】領域抽出部101は、映像データにおける一部の領域を抽出する。この一部の領域とは、画像中の特定の人や車、建物など（あるいはその一部分、例えば人の頭、車のボンネット、建物の玄関など）の物体の領域である。映像では、連続するフレームには同じ物体が写っていることが多いが、物体自身の動きや撮影時のカメラの動きが原因で同じ物体に対応する領域がフレーム間で変化することが多い。

【0044】領域抽出部101は、このような注目している物体の動きや変形に対応して各フレームにおける物体（オブジェクト）の領域を抽出するためのものである。具体的な抽出手法としては、全フレームに渡って人手で領域指定する方法や、M. Kass 他, 「Snakes: Active contour models」(International Journal

of Computer Vision, vol. 1, No. 4, pp. 321-331. July, 1988) に記されているようなSnakesと呼ばれる動的輪郭モデルにより物体の輪郭抽出を連続して行う方法、金子他「ロバスト推定を用いたハイパーメディアコンテンツ作成のための高速移動物体追跡法」(情報処理学会技術報告 CVIM113-1, 1998) に記されているようなブロックマッチングにより求めた物体中の部分領域の移動先から物体全体の変形・移動を推定していく方法、画像解析ハンドブック(第II部第2章, 東京大学出版会, 1991) に記されているような領域成長・分割により似た色を持つ領域を特定する方法などを用いることができる。

【0045】領域の図形近似部(以下、領域図形近似部)102は、領域抽出部101で抽出された映像中の物体の領域を予め定められた図形により近似する。図形としては、矩形、円、楕円、多角形など任意の図形としてよい。また、領域の近似方法も、領域に外接する図形に近似する方法、領域に内接する図形とする方法、領域の重心を近似図形の重心とする方法、領域と近似図形の面積比を等しくする方法、などの方法がある。なお、物体の領域を予め定められた図形により近似するのではなく、近似する対象物体毎に図形の種類をユーザが指定できるようにしてもよいし、近似する対象物体毎にその物体の形状等に応じて図形の種類を自動的に選択するようにしてもよい。

【0046】領域の図形近似は、領域抽出部101での抽出結果が入力されるたびに、フレームごとに行う。もしくは、前後数フレームの領域抽出結果を使って図形近似を行っても良い。数フレームの領域抽出結果を利用する場合には、近似図形の大きさや位置などの変化を数フレームの間で平滑化することにより、近似図形の動きや変形をなめらかにしたり、領域抽出の抽出誤差を目立たなくすることができる。なお、近似図形の大きさは、フレームごとに異なって構わない。

【0047】図形代表点抽出部103は、領域図形近似部102の出力である近似図形を表現する代表点を抽出する。どのような点を代表点とするかは、どのような近似図形を用いるかにより異なる。例えば、近似図形が矩形の場合には4つもしくは3つの頂点を代表点とすることができ、近似図形が円の場合には中心と円周上の一点としたり直径の両端点としたりすることができる。また、楕円の場合には楕円の外接矩形の頂点としたり2つの焦点と楕円上の1点(例えば短軸上の1点)としたりすればよい。任意の閉多角形を近似図形とする場合には、各頂点を図形の代表点とする必要がある。

【0048】代表点の抽出は、領域図形近似部102から1フレーム分の近似図形の情報が出力されるたびに、フレーム単位で行われる。また、各代表点は、水平方向の座標軸Xと、垂直方向の座標軸Yと、により表され

る。

【0049】代表点の曲線近似部(以下、代表点曲線近似部)104は、図形代表点抽出部103で抽出された各代表点の位置をそれぞれ時系列で曲線に近似する。この近似曲線は、各代表点について、X座標、Y座標ごとに、フレーム番号fもしくは映像に付与されているタイムスタンプtの関数として表現される。曲線近似の方法としては、直線による近似、スプライン曲線による近似などがある。

【0050】関連情報記憶部105は、映像データ記憶部100に記憶されている映像データに登場する物体に関する関連情報(もしくは、関連情報自体は別の記憶装置(例えば、インターネットやLAN上のサーバであってもよい)に記憶されていて、その関連情報が記憶されているアドレスの情報)が記憶されている。関連情報は文字、音声、静止画、動画、あるいはそれらを適宜組み合わせたものであってもよいし、プログラムもしくは計算機の動作を記述したデータであってもよい。関連情報記憶部105は映像データ記憶部100と同様に例えばハードディスクや光ディスク、半導体メモリなどで構成される。

【0051】領域データ記憶部106は、代表点曲線近似部104の出力である代表点の時系列的な軌跡を近似した曲線式を表現するデータを含む物体領域データが記憶される記憶媒体である。関数で表現された領域に対応する物体に関する関連情報が関連情報記憶部105に記憶されている場合には、物体領域データには関連情報そのものや関連情報の記録されているアドレスを併せて記録することができる(関連情報記憶部105に関連情報が記録されているアドレスの情報が記憶されている場合には、当該アドレス情報を併せて記録することができる)。領域データ記憶部106も映像データ記憶部100等と同様に例えばハードディスクや光ディスク、半導体メモリなどで構成される。

【0052】なお、映像データ記憶部100、関連情報記憶部105、領域データ記憶部106は、別々の記憶装置によって構成されていてもよいが、それらの全部または一部が同一の記憶装置によって構成されていてもよい。

【0053】また、本物体領域データ変換装置は、計算機上でソフトウェアを実行する形で実現することもできる。

【0054】次に、より具体的な例を用いながら本物体領域データ変換装置の動作について説明する。

【0055】図2は、領域抽出部101による物体の領域を抽出する処理から、領域図形近似部102による領域を図形で近似する処理、図形代表点抽出部103による図形の代表点を抽出する処理、代表点曲線近似部104による代表点を曲線で近似する処理までの一連の処理の概要をより具体的に説明するための図である。

17

【0056】図2では、領域図形近似部102として領域の楕円による近似方法を用い、図形代表点抽出部103として楕円の2つの焦点と楕円上の1点を抽出する方法を用い、代表点曲線近似部104としてスプライン関数による近似方法を用いた場合を例としている。

【0057】図2(a)において、200は処理対象となっている映像中の1フレームを示している。

【0058】201は抽出対象となっている物体の領域を示している。この物体の領域201を抽出する処理は領域抽出部101において行われる。

【0059】202は物体の領域を楕円で近似したものである。物体領域201から楕円202を求める処理は領域図形近似部102において行われる。

【0060】ここで、図3に、物体の領域が平行四辺形で表される場合に、近似楕円を求める方法の一例を示す。図3における点A、B、C、Dが物体の領域である平行四辺形の各頂点である。この場合、まず、辺ABと辺BCのどちらが長いかを計算する。そして、長い方の辺およびその対辺を辺の一部とする最小の長方形を求める。図3の例の場合は点A、B'、C、D'を4頂点とする長方形となる。近似楕円は、例えば、この長方形に内接する楕円と相似で、かつ、点A、B'、C、D'を通る外接楕円とする。

【0061】次に、図2(b)において、203は楕円を表現する図形代表点であり、具体的には2つの焦点および1つの楕円上の点(図2(b)の例では短軸上の1点)である。楕円の焦点は2つの軸上の点や楕円の外接矩形から簡単に求めることができる。以下、一例として、図4における長軸上の2点P0、P1および短軸上の1点Hから焦点FとGを求める方法を説明する。

【0062】まず、長軸と短軸のパラメータであるa、bと、楕円の中心Cと、扁平率eとを、

$$E(P0, P1) = 2 \times a,$$

$$C = (P0 + P1) / 2,$$

$$E(C, H) = b,$$

$$e = (1/a) \times \sqrt{a^2 - b^2}$$

により求める。ここで、E(P、Q)は点Pと点Qのユークリッド距離である。

【0063】このようにして求めたパラメータから、焦点FとGは、

$$F = C + e \times (P0 - C),$$

$$G = C - e \times (P0 - C)$$

により求めることができる。

【0064】このようにして、楕円の代表点F、GおよびHは決定されるが、これらの点を別のフレームにおいて取り出された楕円の代表点と結びつける際にはあいまいさが生じる。すなわち、抽出された2つの焦点と1つ前のフレームにおける2つの焦点とを結びつける組み合わせは2通り存在する。また、短軸と楕円との交点は2つあるため、1つ前のフレームにおいて抽出された楕円

18

上の一点と対応する交点がどちらなのかわからない。そこで、これらを決定する方法について説明する。

【0065】1フレーム前に抽出された2つの焦点をFp、Gpとする。Fpに対応するのがFであるのかGであるのかを判断するために、 $E((Gp - Fp) / 2, (G - F) / 2)$ と $E((Gp - Fp) / 2, (F - G) / 2)$ とを比較する。前者の方が小さい場合には、FpはFに対応させ、GpはGに対応させる。一方、後者の方が小さい場合には、その逆に、FpはGに対応させ、GpはFに対応させる。

【0066】また、1つ前のフレームにおける短軸と楕円との交点をHpとし、現フレームの短軸と楕円との2つの交点をH、H'とする。Hpと対応付ける点としてHとH'のどちらを選択するかは2つの距離、 $E(Hp - (Gp + Fp) / 2, H - (F + G) / 2)$ と $E(Hp - (Gp + Fp) / 2, H' - (F + G) / 2)$ とを算出することにより決定する。前者が小さい場合にはHを選択し、そうでない場合にはH'を選択する。なお、はじめのフレームにおける短軸と楕円との交点Hは2つのうちのどちらを選択してもよい。

【0067】このように楕円から代表点を抽出する処理は図形代表点抽出部103において行われる。

【0068】以上の処理によって取り出された代表点は、映像中の注目物体の移動や撮影カメラの移動により、連続するフレームにおいて位置が異なるのが普通である。そこで、対応する楕円の代表点を時系列に並べ、X座標、Y座標ごとにスプライン関数により近似を行う。本実施形態では、楕円の代表点であるF、G、Hの3点(図4参照)それぞれについてX、Y座標のスプライン関数が必要になるので、合計6つのスプライン関数が生成される。

【0069】以上のスプライン関数による曲線近似は代表点曲線近似部104において行われる。

【0070】なお、代表点曲線近似部104による処理は、当該物体の領域に関する各フレームの代表点の座標値が得られる毎に行う方法(例えば各フレームの代表点の座標値が得られる毎に近似を行うとともに近似誤差を求め、近似誤差が一定の範囲に収まるように近似区間を適宜分割する方法)や、当該物体の領域に関する全てのフレームの代表点の座標値が得られた後に行う方法などがある。

【0071】図2(c)の204は近似されたスプライン関数を3次元的に表現したものである。図2(d)の205は代表点曲線近似部104の出力であるスプライン関数の一例である(1つの代表点の1つの座標軸についてのみ示している)。この例は、近似区間が $t=0 \sim 5$ と $t=5 \sim 16$ の2つに分割された場合(節点が3つとなった場合)を示している。

【0072】このようにして得られたスプライン関数は予め決めておいたデータ形式に従って領域データ記憶部

106に記録される。

【0073】以上のように本実施形態では、映像中の物体の領域を、その近似図形の代表点の時系列的な軌跡

(フレーム番号あるいはタイムスタンプを変数とする代表点の座標の軌跡)を近似した曲線のパラメータとして記述することができる。

【0074】本実施形態によれば、映像中の物体の領域を関数のパラメータのみによって表現できるため、データ量が少なく、ハンドリングの容易な物体領域データを生成することができる。また、近似図形からの代表点の抽出や、近似曲線のパラメータの生成も容易に行うことができる。また、近似曲線のパラメータから近似図形を生成することも容易に行うことができる。

【0075】また、この近似図形として基本的な図形、例えば一つまたは複数の楕円を用い、例えば楕円を二つの焦点と他の1点により代表させれば、ユーザにより指定された任意の座標が物体の領域(近似図形)内か否か(物体の領域を指示しているか否か)を簡単な判定式により判定することができる。さらに、これによって、ユーザによる映像中の動く物体の指示をより容易にすることが

【0076】以下では、領域データ記憶部106に格納される物体領域データのデータ形式について説明する。なお、ここでは、代表点をスプライン関数により近似する場合を例にとって説明するが、もちろん、代表点を他の関数により近似する場合も同様である。

【0077】図5に、映像中の物体の領域を表すスプライン関数と、物体に関連付けられた関連情報とを記録するための物体領域データのデータ形式の一例を示す。

【0078】ID番号400は、物体ごとに付与される識別番号である。なお、このデータは省略されてもよい。

【0079】先頭フレーム番号401と最終フレーム番号402は、当該ID番号の物体の存在を定義する最初と最後のフレーム番号であり、具体的には、映像中で物体が登場し消えるまでのフレーム番号である。ただし、現実映像中で物体が登場し消えるまでのフレーム番号ではなく、例えば映像中で物体が登場したときより後の任意のフレーム番号を先頭フレーム番号としてもよいし、また先頭フレーム番号以降でかつ映像中で物体が消えたときより前の任意のフレーム番号を最終フレーム番号としてもよい。なお、先頭/最終フレーム番号は先頭/最終タイムスタンプで代用することもできる。また、最終フレーム番号402は、物体存在フレーム数または物体存在時間に置き換えてもよい。

【0080】関連情報へのポインタ(以下、関連情報ポインタとも言う)403は、当該ID番号の物体に関連付けられた関連情報データの記録してあるデータ領域のアドレスなどである。関連情報へのポインタ403を使うことにより、物体に関する関連情報を検索したり表示

したりすることが容易にできる。また、関連情報へのポインタ403はプログラムや計算機の動作を記述したデータへのポインタであってもよい。この場合には、当該物体がユーザにより指定されると、計算機が所望の動作を行うことになる。

【0081】なお、関連情報へのポインタ403は、例えば、物体により異なる動作をさせる必要のない場合には、省略することができる。

【0082】なお、以下の説明では、物体領域データ内に関連情報へのポインタを記述する場合について説明するが、物体領域データ内に関連情報そのものを記述するようにしてもよい。また、物体領域データ内に関連情報へのポインタと関連情報そのものを任意に記述可能としてもよい。この場合には、物体領域データ内に記述されているのが関連情報へのポインタか関連情報そのものを示すフラグを付加すればよい。

【0083】近似図形数404は、物体の領域を近似している図形の数である。図2の例においては、一つの楕円で物体領域を近似しているため、図形数は1となる。

【0084】近似図形データ405は、近似図形を表現するための図形代表点の軌跡データ(例えばスプライン関数のパラメータ等)である。なお、近似図形データ405は、近似図形数404に相当する数だけ存在する(近似図形数404が2以上の場合については後述する)。また、物体領域データにおいて、近似図形数404は常に1つとし(従って近似図形データ405も常に1つとなる)、近似図形数404のフィールドを省くようにしてもよい。

【0085】次に、図6に、近似図形データ(405)のデータ構造の一例を示す。

【0086】図形種ID1300は、近似図形としてどのような図形を用いているかを示すためのデータであり、円、楕円、矩形、多角形などを特定する。

【0087】代表点数1301は、図形種IDで特定される図形を代表する点の数を幾つ必要とするかを表す。

【0088】1組の代表点軌跡データ1302、1303は、図形の代表点の軌跡を表現するスプライン関数に関するデータ領域である。一つの図形代表点につき、X座標、Y座標のスプライン関数データが必要になる。従って、スプライン関数を特定する代表点軌跡データは、代表点数(M)×2だけ存在する。

【0089】なお、使用する近似図形の種数を予め1種類、例えば楕円、に限定することも可能である。この場合には、図6の図形種IDのフィールドを省くことも可能である。

【0090】また、図形種IDによって代表点数が一意に特定される場合には、代表点数を省くことも可能である。

【0091】次に、図7に、代表点軌跡データ(130

21

2, 1303) のデータ構造の一例を示す。

【0092】接点フレーム番号1400は、スプライン関数の節点を表しており、この節点まで多項式のデータ1403が有効であることを示している。多項式の係数データの数は、スプライン関数の最高次数により変化する（最高次数をKとすると、係数データの数はK+1となる）。そのため、多項式次数1401を参照する。多項式次数1401の後には、多項式次数+1個に相当する数の多項式係数1402が続く。

【0093】また、スプライン関数は節点間で別の多項式で表現されるため、接点の数に対応した数の多項式が必要になる。従って、節点フレーム番号、多項式の係数などを含むデータ1403は、複数繰り返し記述される。節点フレーム番号が最終フレームと等しくなった場合には、それが最後の多項式係数データであることを意味しているので、代表点軌跡データが終わることがわかる。

【0094】ここで、近似図形として楕円以外の図形を用いた場合について説明する。

【0095】図8は、近似図形として平行四辺形を用いた場合の代表点を説明するための図である。点A、B、C、Dが平行四辺形の頂点である。これらのうち3点が決まれば残りの1点も一意に決まるため、代表点としては4頂点のうちの3頂点とすれば十分である。この例では、A、B、Cの3点を代表点としている。

【0096】また、図9は、近似図形として多角形を用いた場合の代表点を説明するための図である。多角形の場合には、頂点の順序を外周に沿った順にしておく。図9の例では、10個の頂点を持つ多角形なので、N1～N10までの全ての頂点を代表点とする。なお、この場合に、内角が180度未満の頂点のみを代表点とするなどして、頂点数を削減するようにしてもよい。

【0097】ところで、前述したように、これまで説明してきた処理は、計算機上でソフトウェアにより行うこともできる。図10は本実施形態の情報処理装置の処理の流れの一例を表したフローチャートであり、本実施形態の情報処理装置をソフトウェアにより実現する場合には例えば図10のフローチャートに従ったプログラムを作成すればよい。

【0098】ステップS11では、映像データから1フレーム分の映像を取り出す。

【0099】ステップS12では、映像中の特定の物体の領域を抽出する。抽出手法は領域抽出部101と同様の手法を用いることができる。

【0100】ステップS13では、ステップS12の処理結果である領域情報に対して、近似図形による当てはめを行う。これも領域の図形近似部102と同様の手法を用いることができる。

【0101】ステップS14では、ステップS13で近似された図形の代表点を抽出する。

22

【0102】ステップS15では、連続するフレームにおける近似図形の代表点列の位置の曲線近似を行う。

【0103】ステップS16は分岐処理で、今処理した画像が最後のフレームか、または処理したフレームにおいて抽出対象の物体が画像中から消失してしまった（もしくは消失したとみなす）場合にはステップS17を、いずれでもない場合にはステップS11の処理へと分岐する。

【0104】ステップS17では、ステップS15で計算された近似曲線を所定のフォーマットに従って物体領域データとして記録媒体に記録する。

【0105】さて、これまでは一つの物体に対して一つの図形を割り当てて、物体の領域を大まかに表す例を中心に説明をしてきたが、一つの物体の領域を複数の図形により近似し、近似精度を向上させることも可能である。図11は、一つの物体を複数の図形で近似した例である。この例では画像中の人の領域を600から605までの6つの楕円により表している。

【0106】一つの物体を図11のように複数の図形で表す場合には、物体を複数の領域に分割する処理が必要となる。この処理にはどのような方法が用いられていても良いが、例えば人手で直接入力する方法がある。この場合、例えば、マウス等のポインティングデバイスを用いて、画像上で領域を矩形や楕円で囲む、あるいはポインティングデバイスの軌跡により領域を指定する、などの操作で実現することができる。また、人手ではなく自動で行う場合には、例えば、物体の動きのクラスタリングで実現する方法がある。これは、連続するフレーム間で物体中の各領域がどのような動きをしたかを相関法（例えば画像解析ハンドブック、第II部第3章、東京大学出版会、1991を参照）や勾配法（例えば、Determining optical flow, B. K. P. Horn and B. G. Schunck, Artificial Intelligence, Vol. 17, pp. 185-203, 1981を参照）などにより求め、これらの動きの似ているものだけをまとめて領域を形成する方法である。

【0107】このようにして分割された各領域に対しては、それぞれについて図1の構成例や図10の手順例により説明される処理を施すことにより、近似図形データを生成することが可能である。この場合、一つの物体の物体領域データに記述すべきスプライン関数が近似図形の増加に応じて増えることになるため、図12で表されるように近似図形データ405を、近似図形数404に相当する数（この場合、複数）だけ含むデータ構造となる。

【0108】なお、前述したように物体領域データにおいて近似図形数を常に1つとし（従って近似図形データも常に1つとなる）、近似図形数のフィールドを省くようにする場合において、一つの物体を複数の図形で表す

ためには、一つの物体を近似する各図形ごとに、物体領域データを生成すればよい（それらは同じID番号を有することになる）。

【0109】なお、本実施形態では、一つの物体を複数の図形で表す場合に、同一の図形を用いるものとしているが、複数種類の図形を混在して使用可能としてもよい。

【0110】次に、本実施形態の利用方法のバリエーションについて説明する。映像中の物体としては、通常、人や動物、建物や植物などが考えられるが、映像中のどのようなものに対しても本実施形態の処理を行うことができる。例えば、テロップを映像中の物体として扱うことができるので、以下では物体のバリエーションとしてテロップを対象とした場合の処理について説明する。

【0111】テロップとは映像中に付帯される文字情報を指す。米国ではクロズドキャプションと呼ばれる文字情報を付帯させることが義務付けられており、日本の放送においてもテロップを多用する傾向にある。表示されるテロップには、静止しているものもあれば、画面下から上へスクロールするもの、画面右から左へスクロールするものなど何らかの動きを伴うものもある。テロップの表示領域を図形近似し、テロップ文字列を関連情報として記憶しておくことにより、映像の内容を簡単に把握したり、所望の映像を検索することが容易に実現できる。

【0112】領域抽出部101の処理では、人手によりテロップ領域を指定する方法、堀「テロップ認識のための映像からの文字部抽出法」（情報処理学会技術報告99-CVIM-114, pp. 129-136, 1999）に記されているような文字の輝度値およびエッジ情報を用いて文字列抽出法を行う方法、片山他「ニュース映像のテロップ文字認識に基づく記事分類の精度向上」（画像の認識・理解シンポジウム（MIRU'98）講演論文集、Vol. 1, pp. 105-110）に記されているようなエッジの強さを調べて背景とテロップを分離する方法などを用いて、テロップ領域の抽出を行う。得られたテロップ領域からは、各文字あるいは各文字列を切り出すこともできる。また、連続したフレームでテロップ領域内のエッジ情報を比較し、テロップの出現フレームおよび消滅フレームを検出する。

【0113】領域図形近似部102の処理では、領域抽出部101で抽出されたテロップ領域を矩形で近似する。テロップが出現したフレームの番号を物体領域データの先頭フレーム番号（図4もしくは図12の401）に、消滅したフレームを最終フレーム番号（402）に格納する。関連情報へのポインタ（403）にはテロップの文字列情報へのポインタを格納し、近似図形データ（405）としてテロップを囲む矩形領域情報を保持する。複数行のテロップの各行を別々の領域とする場合や各文字を別々の領域とする場合には、行数または文字数

を近似図形数（404）に格納し、それぞれの行あるいは文字を囲む矩形領域データすなわち近似図形データ（405）を該当数だけ格納する。

【0114】図形代表点抽出部103および代表点曲線近似部104での処理は、テロップに特化したものではなく、前述の通りである。

【0115】関連情報記憶部105には、登場するテロップの文字列情報を保存しておき、その情報へのポインタをテロップ領域データ（物体領域データ）が保持する。

【0116】以上により、キーワードを入力し、それに合致あるいは関連する文字列がテロップの文字列情報の中にあれば、その文字列が出現するフレームや時刻を容易に知ることができ、ニュース番組であれば、興味のある記事を検索し、その部分だけを鑑賞することも可能となる。

【0117】なお、この場合には、テロップの文字列情報に、そのフレームや時刻に対応する物体領域データへのポインタを付加するようにすれば、検索をより容易にすることができる。

【0118】以上、テロップを対象とした場合の処理について説明したが、本実施形態の利用方法やどのようなものを対象物体とするかについては、他にも様々なバリエーションが考えられる。

【0119】さて、図2の例では、楕円による近似方法を一例として説明を行ったが、以下では、近似方法の他の例として、矩形による近似方法について説明する。

【0120】図13は、図2と同じ形式の図であるが、ここでは、領域図形近似部102として領域の矩形による近似方法を用い、図形代表抽出部103として矩形の4つの頂点を抽出する方法を用い、代表点曲線近似部104としてスプライン関数による近似方法を用いた場合を例としている。

【0121】図13（a）において、2800は処理対象となっている映像中の1フレームを示している。

【0122】2801は抽出対象となっている物体の領域を示している。この物体の領域2801を抽出する処理は領域抽出部101において行われる。

【0123】2802は物体の領域を矩形で近似したものである。物体領域2801から矩形2802を求める処理は領域図形近似部102において行われる。

【0124】ここで、例えば図13（a）において矩形2802を求める場合の処理手順の一例を図14に示す。すなわち、フレーム2800のマスク画像を例えばラスタースキャンし（ステップS60）、対象画素が物体領域内であるときに（ステップS61）、X座標とY座標のそれぞれについて、それまで保存されている最小値より小さい場合には最小値を更新し、それまで保存されている最大値より大きい場合には最大値を更新する（ステップS62）ことを、全ての画素について繰り返

25

しチェックすることによって、X座標とY座標のそれぞれについて物体領域を示す画素位置の最小値および最大値を求めれば、矩形2802の4つの頂点座標を得ることができる。

【0125】このような方法は処理が簡易な点で優れているが、例えば図15のように細長い物体3001が画面3000に対して斜めの姿勢で存在しているときには、近似矩形3002内には非物体領域が特に多く含まれてしまう。また、細長い物体物体が回転すると矩形2802の大きさ、形状が変化する。これらは、物体指定の際の弊害となる場合がある。

【0126】そこで、矩形の大きさができるだけ小さくなり（近似矩形内の非物体領域ができるだけ少なくなり）、対象物体の姿勢も反映させることのできる近似方法の一例を示す。

【0127】図16(a)において、3100は処理対象となっている映像中の1フレームを示している。

【0128】3101は抽出対象となっている物体の領域を示している。この物体の領域3101を抽出する処理は領域抽出部101において行われる。

【0129】3102は物体の領域を矩形で近似したものである。この近似矩形は図13(a)の矩形2802とは異なり、傾きを持っている。矩形内の非物体領域も少なく、対象が回転してもその形状は一定である。物体領域3101から矩形3102を求める処理は領域図形近似部3102において行われる。

【0130】図17に、この場合の処理手順の一例を示す。この処理手順例は、対象物体領域の慣性主軸を求め、これに基づいて近似図形を求めるようにしたものである。

【0131】図16(b)において、3103は、対象物体領域の重心を示している。

【0132】3104は、対象物体領域の慣性主軸を示している。3105は、3104に垂直な直線である。

【0133】まず、対象物体領域の慣性モーメント m_{20} 、 m_{02} 、 m_{11} を求める（ステップS70～S72）。

【0134】マスク画像を $f(x, y)$ とすると、 $f(x, y)$ は領域内では1で、領域外では0である。対象領域の慣性モーメントは、

$$m_{ij} = \sum \sum x^i y^j f(x, y)$$
 で表せる。

【0135】ここで、原点を通る直線 $y = x \tan \theta$ についての $f(x, y)$ の慣性モーメントは、

$$m_\theta = \iint (x \sin \theta - y \cos \theta)^2 f(x, y) dx dy$$
 で得られる。 θ を変化させたときに m_θ を最小にする角度を θ_0 とする。一通りの角度しかないとき、直線 $y = x \tan \theta_0$ を慣性主軸と呼ぶ。 $\tan \theta_0$ は、2次方程式、

$$\tan^2 \theta + \{ (m_{20} - m_{02}) / m_{11} \} \tan \theta - 1 = 0$$

26

0

の解として求まる。これにより重心3103の周りで $\tan \theta_0$ を求めると、対象物体の慣性主軸が得られる（ステップS73）。

【0136】次に、慣性主軸に平行で物体領域に外接する直線と、慣性主軸に垂直で物体領域に外接する直線を求める（ステップS74）。図16(b)において、直線3106、3107は、慣性主軸3104に並行な直線であり、対象物体領域に外接する。直線3108、3109は、直線3105に平行な直線であり、対象物体領域に外接する。矩形3102は、直線3106、3107、3108、3109によって形成される（ステップS75）。

【0137】なお、対象物体が円の場合には慣性主軸が求まらないが、このような場合には例えば図14で説明した手順で近似矩形を求めればよい。

【0138】ところで、矩形より楕円で物体領域を表現する方が適当な場合もある。図18に、物体の領域が矩形で表される場合に、その矩形から近似楕円を求める方法の一例を示す。図19に、この場合の処理手順の一例を示す。

【0139】図18において、対象物体領域3300と外接矩形3301が得られているものとする。

【0140】まず、近似矩形の内接楕円および外接楕円を求める（ステップS80）。

【0141】図18において、楕円3302は矩形3301の内接楕円であり、楕円3303は矩形3301の外接楕円である。

【0142】次に、内接楕円3302の大きさを少しずつ外接楕円3303に近づけていき（ステップS81）、物体領域を全て包含する楕円3304を求め（ステップS82）、近似楕円とする。なお、繰り返し処理において、一回に内接楕円3302の大きさを拡大する単位は、予め定めておいてもよいし、内接楕円3302の大きさと外接楕円3303の大きさの差分に応じて決定してもよい。

【0143】また、上記とは逆に、外接楕円3303の大きさを内接楕円3302に近づけていってもよい。この場合には、外接楕円3303は最初から物体領域を全て包含しているので、例えば、繰り返し処理において、始めて物体領域に包含されない部分を生じた楕円の、一回前における楕円を、近似楕円とすればよい。

【0144】次に、図形代表点抽出部103では、近似矩形や近似楕円の代表点を求める。矩形の代表点は4つもしくは3つの頂点を代表点とすることができ、楕円の場合は楕円の外接矩形の頂点としたり、2つの焦点と楕円上の一点としたりすることができる。

【0145】次に、代表点曲線近似部104では、時系列的に得られている代表点の軌跡をスプライン関数などで近似する。このとき、時系列同士の対応付けが重要で

27

ある。例えば、近似図形が矩形で代表点を頂点とする場合には、隣接するフレーム間で頂点同士の対応付けを行う必要がある。

【0146】図20に、対応付け方法の一例を示す。また、図21に、対応付け処理の手順の一例を示す。

【0147】図20において、3500は近似矩形の重心である。前のフレームにおける矩形3501と現在のフレームにおける矩形3502が得られている。矩形3501と3502のいずれかを平行移動させ、重心位置を一致させる（重心位置を一致させた状態が図20である）。次に、それぞれの矩形の頂点同士の距離 $d_1 \sim d_4$ を計算し、全ての頂点の組み合わせで距離の和を求める（ステップS90、S91）。距離の和が最小となる組み合わせを求め、対応付けする（ステップS92）。

【0148】近似図形から代表点を得る際に、一定の規則で代表点を取得しておく、組み合わせの数を削減することが可能である。例えば、矩形では時計周りに頂点座標を代表点として保存するようにすると、対応付けの組み合わせは4通りで済む。

【0149】なお、この方法では対応付けが難しい場合がある。例えば、近似矩形が隣接フレームにおいて正方形に近い形状であり且つ45度回転移動している場合には、対応付けが難しい（2通りの組み合わせにおいて距離の和が同じような値になる）。そこで、このような場合には、例えば、近似矩形内の物体領域同士の排他的論理和を取り、その面積が最小となる組み合わせを採用する方法、あるいは物体領域のテクスチャの絶対差分を求め、差分値が最小となる組み合わせを求める方法などを用いればよい。

【0150】続いて、以下では、本発明を適用して物体領域の軌跡を記述する際に、図6、図7で例示した近似データ構造とは別の構造を用いる例について説明する。

【0151】図22は、近似図形データおよび物体領域の代表点軌跡データの記述フォーマットの他の例である。なお、図22では、代表点軌跡データの部分については、1つの代表点についてのみ示してある（実際には、代表点の個数に対応して記述される）。

【0152】ここでは、多項式の最高次数を2次として説明する。

【0153】前述した例（図5、図6、図7）では、多項式スプライン関数の全ての係数を記述していたのに対して、ここでの記述方法では、スプライン関数の節点の座標と、スプライン関数の2次の係数に関連する値との組み合わせにより記述する。この記述方法の利点は、節点が容易に取り出せるため、大まかな物体の軌跡が簡単にわかるという点である。

【0154】以下、この記述方法について詳細に説明する。

【0155】図22中、図形種ID（3900）は、物体の形の近似に用いた図形の種類を特定する。例えば、

28

物体の重心のみ（CENTROID）、矩形（RECTANGLE）、楕円（ELLIPSE）や、それらの組み合わせを指定できる。図23は、図形の種類と図形種IDの割り当て例である。また、代表点数3901は、図形の種類によって定まる代表点軌跡の数を表す。

【0156】節点数（N）3902は、代表点軌跡を表すスプライン関数の節点の数を表す。各節点に対応するフレームは、時間として表され、節点時刻3903に格納される。節点時刻は、設定数だけあるため、配列として記述しておく（3904）。

【0157】同様に、各節点のx座標、y座標もそれぞれ節点X（3905）および節点Y（3907）の配列（3906、3908）として記述される。

【0158】一次関数フラグ3909は、節点間のスプライン関数として一次関数だけが用いられているかどうかを表す。一部分でも2以上の多項式を用いる場合には、このフラグはオフにしておく。このフラグを用いることにより、近似関数として一次関数のみに使われる場合に以下で説明する関数特定情報3910を1つも記述しなくて済むため、データ量を削減できるというメリットがある。なお、必ずしもこのフラグは必要ではない。

【0159】関数特定情報に含まれる関数ID（3911）、関数パラメータ（3912）はそれぞれ多項式スプライン関数の次数と、その係数を特定するための情報を表す。図24に、それらの一例を示す。ここで、 t_a 、 t_b は連続する節点の時刻、 $f(t)$ は $[t_a, t_b]$ の区間のスプライン関数、 f_a 、 f_b は時刻 t_a 、 t_b における節点の座標を表している。1次多項式を用いるときは節点のみの情報で十分なので、関数パラメータは記述されていないが、2次多項式の場合には係数を特定するための情報として一つの値が関数パラメータに記述される。なお、図24の例では、2次の係数が用いられているが、例えば、二次曲線上の f_a 、 f_b 以外の1点など、他の値を用いることもできる。

【0160】本記述方法では、節点の情報と関数パラメータの情報により、図24の制約条件を用いて全ての区間におけるスプライン関数が再現できる。

【0161】関数特定情報は、（節点数-1）に相当する個数のものが存在し、これらは配列となって記述される（3913）。

【0162】なお、上記では、多項式の最高次数を2次として説明したが、もちろん、多項式の最高次数を3次以上とすることも可能である。

【0163】続いて、以下では、関連情報のバリエーションを示す。

【0164】図25は、主として監視映像に用いられる関連情報のデータ構造の一例を示す。実際のデータはこれらの項目のうち、少なくとも一つを含めればよい。オブジェクト種4201は、“車両”や“人”など、近似されている物体の種類を示すデータである。識別情報4

202は、“人名”や“車両のナンバー”、“車種”など、実際の物体を識別するためのデータである。動作内容4203は、“歩く”や“走る”などの物体の動作を示すデータである。

【0165】図26は、主として商用コンテンツやハイパーメディア・コンテンツなどに用いられる関連情報のデータ構造の一例を示す。実際のデータはこれらの項目のうち、少なくとも一つを含めればよい。名前4301は、物体の名前を示すデータである。物体が映画などに登場する人物である場合は、役者や俳優の名前を特定する。著作権情報4302は、物体の著作権者などの著作権に関わる情報を示すデータである。コピー許可情報4303は、物体を近似する図形に含まれる範囲の映像情報を切り出して再利用することを許可するか否かを示すデータである。フットマーク4304は、物体が最後に編集された日時を示すデータである。関連情報のURL4305は、ハイパーリンクを用いて、物体の関連情報を表示するときに参照すべきデータをURLによって表現したものである。アクセス制限情報4306は、物体ごとに視聴やハイパーリンクによるジャンプの許可/不許可の情報や許可条件を設定するためのデータである。課金情報4307は、物体ごとの課金情報を示すデータである。注釈データ4308は、物体の関連情報や操作の手助けとなるデータである。

【0166】なお、図25や図26の関連情報は、物体領域データ内に記述すると好ましい。

【0167】次に、映像データや物体領域データの提供方法について説明する。

【0168】本実施形態の処理により作成された物体領域データがユーザの用に供される場合には、作成者側からユーザ側に何らかの方法で物体領域データを提供することが必要である。この提供の方法としても以下に例示するように種々の形態が考えられる。

(1) 映像データとその物体領域データとその関連情報とを1つ（または複数の）記録媒体に記録して同時に提供する形態

(2) 映像データとその物体領域データとを1つ（または複数の）記録媒体に記録して同時に提供するが、関連情報は別途提供するかもしれない（後者は例えば提供しなくてもユーザがネットワーク経由等で別途取得できる場合）形態

(3) 映像データを単独で提供し、別途、物体領域データと関連情報とを1つ（または複数の）記録媒体に記録して同時に提供する形態

(4) 映像データ、物体領域データ、関連情報を別々に提供する形態

上記は主に記録媒体により提供する場合であるが、その他にも、一部または全部を通信媒体で提供する形態も考えられる。

【0169】（第2の実施形態）第1の実施形態では映

像中の物体の領域を近似する図形の代表点を抽出し、物体領域データへの変換を行ったが、第2の実施形態は、映像中の物体領域中の特徴点を抽出し、物体領域データへ変換するようにしたものである。

【0170】本実施形態では、第1の実施形態と相違する点を中心に説明する。

【0171】図27に、本実施形態に係る物体領域データ変換装置の構成例を示す。図27に示されるように、本物体領域データ変換装置は、映像データ記憶部231、特徴点抽出部233、特徴点の曲線近似部234、関連情報記憶部235、領域データ記憶部236を備えている。

【0172】図27において、映像データ記憶部230は第1の実施形態の映像データ記憶部100と、関連情報記憶部235の第1の実施形態の関連情報記憶部105と、領域データ記憶部236は第1の実施形態の領域データ記憶部106とそれぞれ同じ機能を持つものである。

【0173】特徴点抽出部233は、映像中の物体領域から特徴的な点を1つ以上取り出す。特徴点としては、種々のものが考えられるが、例えば、物体の角（例えば、Gray-level corner detection, L. Kitchen and A. Rosenfeld, Pattern Recognition Letters, No. 1, pp. 95-102, 1982による方法など）、物体の重心などが考えられる。

【0174】特徴点の曲線近似部（以下、特徴点曲線近似部）234は、基本的には第1の実施形態における代表点曲線近似部104と同様の機能を持つ。すなわち、特徴点抽出部233で抽出された特徴点の位置を時系列で曲線に近似する。この近似曲線は、X座標、Y座標ごとにフレーム番号fもしくは映像に付与されているタイムスタンプtの関数として表現され、直線近似やスプライン曲線による近似などで曲線に近似される。近似後のデータの記録方法は第1の実施形態と同様である。

【0175】また、本実施形態における物体領域データは基本的には第1の実施形態の物体領域データ（図5）と同様である。ただし、近似図形数のフィールドは不要である。また、「近似図形データ」が「特徴点データ」となる。

【0176】物体領域データの中の特徴点データも基本的には第1の実施形態の近似図形データ（図6）と同様である。ただし、「代表点数」が「特徴点数」となり、「代表点軌跡データ」が「特徴点軌跡データ」となる。また、図形種IDは不要である。

【0177】特徴点データの中の特徴点軌跡データは第1の実施形態の代表点軌跡データ（図7）と同様である。

【0178】図28に本実施形態の物体領域データ変換装置の処理の流れの一例を表したフローチャートを示

31

す。全体的な流れとしては第1の実施形態と同様で、図10のステップS12～S14の部分の特徴点の抽出とし、ステップS15の部分の代表点の特徴点としたものである。

【0179】もちろん、本実施形態もソフトウェアによっても実現可能である。

【0180】以上のように本実施形態では、映像中の物体の領域を、その特徴点の時系列的な軌跡（フレーム番号あるいはタイムスタンプを変数とする特徴点の座標の軌跡）を近似した曲線のパラメータとして記述することが

【0181】本実施形態によれば、映像中の物体の領域を関数のパラメータのみによって表現できるため、データ量が少なく、ハンドリングの容易な物体領域データを生成することができる。また、特徴点の抽出や、近似曲線のパラメータの生成も容易に行うことができる。

【0182】また、本実施形態によれば、ユーザにより指定された任意の座標が物体の領域を指示しているか否かを極めて簡単に判定することができる。さらに、これによって、ユーザによる映像中の動く物体の指示をより

【0183】ところで、第1の実施形態の物体の領域の近似図形の代表点をもとにした物体領域データと、第2の実施形態の物体の領域の特徴点をもとにした物体領域データとを混在して使用可能とすることもできる。この場合には、例えば、物体領域データのフォーマットとしては第1の実施形態のものに、物体の領域の近似図形の代表点をもとにしたものかあるいは物体の領域の特徴点をもとにしたものかを識別するフラグを設ければよい。あるいは、第1の実施形態の物体領域データのフォーマットにおいて、該フラグを設けるのではなく、例えば、図形種IDが特定の値のときに、物体の領域の特徴点をもとにしたものであることを示すものとし、それ以外のときに、物体の領域の近似図形の代表点をもとにしたものであることを示すものとしてもよい。

【0184】さて、これまでは、物体領域データの構造とそれを作成する側について説明してきたが、以下では、このような物体領域データを利用する側について説明する。

【0185】（第3の実施形態）第3の実施形態では、映像中の物体に関連情報を含む物体領域データが付与されている場合に、ユーザが（主にGUI画面上で）物体を指示することにより、関連情報を呈示（文字、静止画、動画等の表示、音声の出力等）させ、あるいは関連するプログラムを実行させるような場合について説明する。

【0186】図29に、本実施形態に係る情報処理装置の構成例を示す。図29に示されるように、本情報処理装置は、映像データ表示部301、制御部302、関連情報呈示部303、指示入力部304を備えている。

32

【0187】映像データ表示部301は、図示しない記録媒体等から入力した映像データを液晶表示装置もしくはCRT等に表示するためのものである。

【0188】指示入力部304は、ユーザがマウス等のポインティングデバイスもしくはキーボードなどを用いて、液晶表示装置もしくはCRT等に表示された映像中の物体を指示するなどの操作を行うことを可能とし、そのユーザからの入力を受け付けるためのものである。

【0189】制御部302は、詳しくは後述するが、例えばユーザが画面上で指示した座標と図示しない記録媒体等から入力した物体領域データとに基づいてユーザが映像中の物体を指示したか否かを判定するとともに、ユーザが映像中のある物体を指示したと判定された場合に、物体領域データの関連情報へのポインタを参照し、その物体の関連情報を取得させ、これを呈示させる制御等を行うためのものである。

【0190】関連情報呈示部303は、制御部302からの指示に応じて、関連情報を（記録媒体やネットワークを介したサーバ等から）取得し、呈示するためのものである。

【0191】なお、関連情報へのポインタがプログラムや計算機の動作を記述したデータへのポインタである場合には、当該計算機が所望の動作を行うことになる。

【0192】もちろん、本実施形態もソフトウェアによっても実現可能である。

【0193】次に、第1の実施形態のように物体の領域が近似図形として表現されている場合に対応する処理手順について説明する。

【0194】図30に、本処理手順の一例を示す。ただし、図30のフローチャートでは、映像の再生中表示されている映像内をマウスカーソル等のポインティングデバイスを用いて指示された際の処理についてのみ示している（基本的には制御部302の処理に相当するものである）。

【0195】ステップS31では、ポインティングデバイス等により指示された画面上の座標が、映像中の画像のどこに相当するかを計算する。さらに、指示された瞬間に再生を行っていた映像のフレーム番号を取得する。なお、フレーム番号ではなく、タイムスタンプを用いてもよい（以下では、フレーム番号として説明する）。

【0196】ステップS32では、映像に付随している映像中の物体の物体領域データから、物体が指示されたフレーム番号において映像中に存在している物体のみを選択する。これは、物体領域データにおける先頭フレーム番号および末尾フレーム番号を参照することにより容易に実行できる。

【0197】次に、ステップS33では、物体領域データから取り出したスプライン関数データ（図6、図7参照）を用いて、物体が指示された際の映像表示フレーム番号における近似図形の代表点の座標を算出する。これ

により、例えば第1の実施形態の例（図2／図4）においては楕円の二つの焦点FとGおよび楕円上の点Hが求まる。

【0198】ステップS34では、上記で求められた代表点と物体領域データの図形種IDに応じて定まる判定手順に基づいて、ポインティングデバイス等により指示された座標が、物体（すなわち、近似図形）の内部か否かを判定する。

【0199】第1の実施形態で説明したように楕円を2つの焦点と1つの楕円上の点により代表させている場合には簡単に判定ができる。

【0200】例えば、第1の実施形態と同様に、E（P、Q）により点Pと点Qのユークリッド距離を表すことにすれば、ポインティングデバイスにより指示された座標Pが楕円内部である場合には、 $E(F, P) + E(G, P) \leq E(F, H) + E(G, H)$ が成り立つ。

【0201】楕円外部である場合には、逆に、 $E(F, P) + E(G, P) > E(F, H) + E(G, H)$ となる。このような不等式を用いて、指示された点が物体内部か外部かを判定し、その判定結果によりステップS35を行うか否かを決定する。

【0202】また、映像中の物体領域の近似図形として平行四辺形を用いている場合には、楕円の場合のような一つの不等式ではなく、4つの不等式を用いて任意の座標が物体の内部かどうかを判定する。

【0203】例えば、図8のA、B、C点を代表点としている場合、点Dを、 $D = C + A - B$ により求める。次に、点Aと点Bを通る直線上の点をQとし、直線式を、 $f_{A,B}(Q) = 0$ で表すことにする。点Pが図形内部にあるときは、2つの不等式、すなわち、

$$f_{A,B}(P) * f_{C,D}(P) < 0, \\ f_{B,C}(P) * f_{D,A}(P) < 0$$

が同時に成り立つ。ただし、

$$f_{A,B}(P) = (y_A - y_B) * (x - x_A) - (x_A - x_B) * (y - y_A)$$

$$f_{B,C}(P) = (y_B - y_C) * (x - x_B) - (x_B - x_C) * (y - y_B)$$

$$f_{C,D}(P) = (y_D - y_C) * (x - x_C) - (x_D - x_C) * (y - y_C)$$

$$f_{D,A}(P) = (y_A - y_D) * (x - x_D) - (x_A - x_D) * (y - y_D)$$

とし、 $P = (x, y)$ 、 $A = (x_A, y_A)$ 、 $B = (x_B, y_B)$ 、 $C = (x_C, y_C)$ 、 $D = (x_D, y_D)$ とした。

【0204】なお、1つの物体が複数の近似図形で近似されている場合には（図5／図12の近似図形数参照）、上記の処理は各近似図形について行われる。

【0205】ステップS35は、指示された点が物体領域内である場合にのみ行われる処理である。この場合、物体領域データに含まれている「関連情報へのポイン

タ」を参照し、このポインタ情報に基づいて関連情報を取得し、その表示等を行う（図29の構成例では、これを関連情報呈示部303により行う）。また、関連情報としてプログラムが指定されている場合には、指定プログラムを実行したり、またその他、指定されている所定の動作を行う。なお、物体領域データに関連情報そのものが記述されている場合にはこれを表示等すればよい。

【0206】図31に、関連情報として映像中の物体の説明が付与されている場合の一例を示す。映像800の再生中にポインティングデバイス802により指示された座標が物体801（を近似した図形の）領域内部であった場合、関連情報803が表示される。

【0207】ステップS36は、分岐処理であり、物体を指示されたフレームにおいてまだ他に物体領域データを持つ物体が存在しているかを判定し、存在している場合にはステップS32へ処理を進め、そして存在しない場合には終了となる。

【0208】次に、第2の実施形態のように物体の領域がその物体の特徴点として表現されている場合に対応する処理手順について説明する。

【0209】ここでは、上記した第1の実施形態のような場合に対応するものとの相違点を中心に説明する。

【0210】図32に、本処理手順の一例を示す。ただし、図32のフローチャートでは、映像の再生中表示されている映像内をマウスカーソル等のポインティングデバイスを用いて指示された際の処理についてのみ示している（基本的には制御部302の処理に相当するものである）。全体的な流れとしては図30のフローチャートと同様であるので、ここでは相違する点を中心に説明する（ステップS41、S42、S45、S46は、それぞれ、ステップS31、S32、S35、S36と同様である）。

【0211】ステップS43では、物体領域データから表示フレーム番号における物体の特徴点の位置座標を算出する。特徴点が複数ある場合には、全ての特徴点の座標を算出する。

【0212】ステップS44では、ステップS43で算出された特徴点の位置とクリックされた座標との間の距離を、全ての特徴点について計算し、予め決められているしきい値よりも距離が小さい特徴点が1つ以上存在するか1つも存在しないかを判定する。あるいは、上記距離をある特徴点について計算し、それを予め決められているしきい値と比較する処理を繰り返し行い、該しきい値よりも距離が小さい特徴点が1つ見つければ、その時点でこの処理をうち切るようにしてもよい。該しきい値よりも距離が小さい特徴点が1つ以上存在する場合にはステップS45へ処理を進め、存在しない場合にはステップS46へと処理を進める。

【0213】このような処理により、物体の領域付近がポインティングデバイス等により指定された場合に、物

35

体の特徴点座標から関連情報の表示を行うことができる。

【0214】（第4の実施形態）次に、第4の実施形態として、関連情報の表示等が可能な物体の領域を物体領域データを用いて明示的に表示する（ユーザに知らしめる）場合について説明する。この場合、関連情報の表示等が可能な物体には、関連情報ポインタを含む物体領域データが予め付与されている必要がある。

【0215】本実施形態のブロック構成は例えば第3の実施形態（図29）と同様である。

【0216】もちろん、本実施形態もソフトウェアによっても実現可能である。

【0217】以下、第1の実施形態のように物体の領域が近似図形として表現されている場合について説明する。

【0218】図33に、本実施形態における処理手順の一例を示す。

【0219】ここでは、近似図形が楕円である場合を例にとって説明する。もちろん、他の近似図形の場合にも同様である。

【0220】ステップS51では、現在表示中の映像のフレーム番号を取得する。なお、フレーム番号ではなく、タイムスタンプを用いてもよい（以下では、フレーム番号として説明する）。

【0221】ステップS52では、ステップS51で取得したフレーム番号において映像中に存在している物体を選択する。これは、映像に付与されている物体領域データの先頭フレーム番号と末尾フレーム番号の間に表示フレーム番号があるデータを見つけることにより行われる。

【0222】次に、ステップS53では、ステップS52で選択された物体の物体領域データからスプライン関数データ（図6、図7参照）を取り出し、表示フレームにおける図形の代表点の座標を算出する。

【0223】ステップS54では、物体領域データの図形種IDを参照してステップS53で算出された代表点で表現される近似図形を求め、該近似図形（例えば楕円領域）内の画像表示を変更する。

【0224】この変更方法には様々な方法が適用可能である。例えば、近似図形が楕円である場合、その楕円領域内の画像の輝度を予め定められた値だけ増加させる。この増加分を ΔY 、表示変更前の輝度を Y 、表示可能な輝度の上限を Y_{max} としたとき、楕円内の画素は $MIN(Y + \Delta Y, Y_{max})$ の輝度で表示を行い、楕円外の画素は輝度 Y で表示を行う。ここで $MIN(a, b)$ は a, b のうちの小さい方の値をとる関数である。

【0225】図34はこのような輝度を増加させる方法で物体領域を表示させた例である（なお、図34においてハッチングされた部分は輝度の変更がなく、ハッチングされた部分は輝度が増加されたことを示している）。 50

36

図34（a）の画面1000は表示中のフレームでステップS54の表示変更処理前の表示状態であり、1001は映像中の物体領域データを持つ物体を表している。

図34（b）の画面1002はステップS54の表示変更後の表示で、1003は物体1001の領域を近似した楕円領域である。楕円領域1003内だけ表示が明るくなり、関連情報の表示等が可能な物体であることがわかるようになっている。

【0226】なお、1つの物体が複数の近似図形で近似されている場合には（図5／図12の近似図数参照）、上記の処理は各近似図形について行われる。

【0227】ステップS55は表示変更を行うべき物体が他にあるかどうかを判定する処理で、表示フレーム番号が物体領域データの先頭フレーム番号と末尾フレーム番号の間となっている未処理の物体があるかどうかを調べる。未処理の物体が存在する場合にはステップS52から処理を繰り返し、存在しない場合には処理を終了する。

【0228】このように、本物体領域データを用いて特定される映像中の物体の領域のうち、関連情報を持つ物体の領域の表示を変更することにより、関連情報があるかないかを一目で確認することが可能になる。

【0229】関連情報の表示等が可能な物体の領域を明示的に示す方法としては、以上説明したような物体の領域内の輝度を変化させる方法の他に、様々な方法を用いることができる。以下、そのいくつかのバリエーションを示す。なお、以下の物体領域データを用いる各処理の手順は、基本的には図33のフローチャートと同様で、ステップS54を各々の該当する処理に変更すればよい。

【0230】図35に例示する表示方法は、画像1600の外に関連情報を持つ物体の位置を表示する方法である。1601および1602が関連情報を持つ物体である。1603および1604はそれぞれ縦軸方向、横軸方向の物体の位置を表示するバーである。関連情報を持つ物体1601に対応する表示が1605および1606である。これらは、1601の領域を縦軸方向、横軸方向に射影した領域に、目印となるバーを表示したものである。同様に、1602に対しては1607、1608が領域を表示するバーである。

【0231】物体の領域の各方向への射影は、これまでの実施形態で説明したようにして物体領域データの近似図形データから求めた当該フレームでの近似図形の代表点の座標と、図形種IDとから、容易に求めることができる。

【0232】なお、異なる物体に対しては、領域を示すバーは異なる表示形態（例えば異なる色）を用いるようにすると好ましい。

【0233】本方法によれば、ユーザは1603と1604に表示されたバーに合わせて画像内をポインティン

グデバイスで指定することにより、関連情報を表示させることができる。

【0234】図36に例示する表示方法は、画像1700の外に関連情報を持つ物体の位置を表示する他の方法である。画像1700の中に関連情報を持つ物体1701および1702が存在している。縦方向と横方向の関連情報を持つ物体位置表示バー1703と1704に物体の位置が表示される。図35の例と異なり、表示バーには物体の重心位置のみが表示される。1705および1706は物体1701の重心位置を表示している。また、1707および1708は物体1702の重心位置を表示している。

【0235】物体の領域の各方向の重心も、これまでの実施形態で説明したようにして物体領域データの近似図形データから求めた当該フレームでの近似図形の代表点の座標と、図形種IDとから、容易に求めることができる。

【0236】本方法によれば、物体の大きさが大きい場合でも、物体位置表示バー上での表示が小さいため、見やすい表示となる。また、多くの物体位置を表示する場合にも同様に見やすい表示となる。

【0237】図37に例示する表示方法は、画像1800の外に関連情報リストを表示する方法である。画像1800に関連情報を持つ物体1801および1802が含まれている。1803は関連情報を持つ物体のリストである。このリストには、現在表示中の画像フレームにおいて、関連情報を持つ物体の情報が表示される。図37の例では、物体の名称が表示されているが、これは、当該フレーム中に存在する物体の物体領域データから関連情報をたどった結果、得られたものである。

【0238】本方法によれば、ユーザは1801または1802の領域をポインティングデバイスで指定するほか、リストに表示された名称を指定しても関連情報を表示させることができる。また、1803のリストに表示されている番号を指定しても関連情報が表示できるように構成可能であるため、リモコンなどポインティングデバイスがない場合でも対応可能である。

【0239】図38に例示する表示方法は、画像1900の中の関連情報を持つ物体1901および1902をアイコンにより表示し、関連情報があることを示す方法である。アイコン1903が物体1901に、そしてアイコン1904が物体1902に対応している。

【0240】上記のアイコンについては、例えば、これまでの説明のようにして近似図形を求め、当該フレームの映像データ中から該求めた近似図形を内包する所定の大きさの矩形の領域を切り出し、該切り出した矩形の領域を適宜縮小するなどして、アイコンとして表示すればよい。

【0241】本方法によれば、映像中の物体の領域を指定するほか、アイコンを直接指定しても関連情報の表示

が行える。

【0242】図39に例示する表示方法は、関連情報を持つ物体の領域を示すマップの表示により、関連情報の存在を示す方法である。画像2000に関連情報を持つ物体1901および1902が含まれている。2003は関連情報を持つ物体領域マップであり、マップは画像2000の中のどこに関連情報を持つ物体領域があるかを示している。2004は物体2001を、2005は物体2002を表している。

【0243】マップ2003は、画像2000を縮小したものであり、物体領域以外の画像表示は行わない（前述のようにして求めた近似図形のみ、縮小画像中の該当する位置に表示する）。

【0244】なお、異なる物体に対しては、領域を示すバーは異なる表示形態（例えば異なる色）を用いるようにすると好ましい。

【0245】本方法によれば、画像2000中の物体を直接指定する他、マップ2003上で表示されている物体領域を指定しても、関連情報の表示等を行うことができる。

【0246】図40に例示する表示方法は、画像中の関連情報を持つ物体をポインティングデバイスで指定することを容易にするために、映像再生をマウスカーソルの位置でコントロールするようにした方法である。2100および2102は表示画面全体を表しており、2101および2103は表示画面上の画像が表示されている領域を表している。図40(a)の表示画面2100においてはマウスカーソル2104が画像2101の領域外にあるので、画像を通常の表示速度で再生させる。一方、図40(b)の表示画面2102ではマウスカーソル2105が画像領域2103の中に入っているため、画像の再生速度を遅くする、もしくは静止させる。

【0247】なお、マウスカーソルが画像領域内に入った場合に、当該フレーム内に関連情報を持つ物体が存在するか否かを調べ（例えば当該フレーム番号と各物体領域データの先頭/最終フレーム番号を比較して調べる）、当該フレーム内に関連情報を持つ物体が存在するときだけ、画像の再生速度を遅くする、もしくは静止させるようにしてもよい。

【0248】例えば映像中で高速に動いている物体をマウスカーソルで指定するのは難しい場合があるが、本方法によれば、このようにマウスカーソルの位置により再生速度を変化させることで、物体指定時の物体の動きを遅くさせもしくは静止させることができ、指定が容易になる。

【0249】図41に例示する表示方法は、画像中の関連情報を持つ物体をポインティングデバイスで指定することを容易にするための他の方法である。2500は再生されている映像である。2501は画像を獲得するポインタであり、このボタンをマウスポインタ2502で指

10

20

30

40

50

39

定すると、指定された時点で表示されていた画像が獲得され、獲得画像表示部 2503 に表示される。2500 は 2501 指定後も再生が続けられる。しかし、獲得画像はしばらく 2503 に表示されるため、2503 内に表示されている物体を指定することにより、指定された物体の関連情報を表示させることができる。

【0250】なお、画像を獲得するボタン 2501 は必ずしも必要ではなく、2501 を設けずにマウスカーソルが映像表示部 2500 内に入ったときに自動的に画像が獲得されるようにしておいても良い。

【0251】また、上記ボタンが指定された場合あるいはマウスカーソルが画像領域内に入った場合に、当該フレーム内に関連情報を持つ物体が存在するか否かを調べ（例えば当該フレーム番号と各物体領域データの先頭／最終フレーム番号を比較して調べる）、当該フレーム内に関連情報を持つ物体が存在するときだけ、画像を獲得し、表示するようにしてもよい。

【0252】本方法によれば、ユーザは静止した画像から容易に関連情報の表示を指定できるようになる。

【0253】以上、いくつかのバリエーションを例示したが、これらの他にも、関連情報の表示等が可能な物体の領域を明示的に示す方法や、指示をしやすくする方法など、ユーザの利用を支援する方法は、様々なものが考えられる。

【0254】次に、第 2 の実施形態のように物体の領域がその物体の特徴点として表現されている場合について説明する。

【0255】ここでは、上記した第 1 の実施形態のような場合に対応するものとの相違点を中心に説明する。

【0256】まず、この場合のフローチャートは、基本的に図 33 のフローチャートと同様で、ステップ S54 の近似図形の代表点の算出が、物体の特徴点の算出となり、ステップ S55 が該特徴点をもとにした種々の該当する処理となる。

【0257】前述した図 34 では、物体に対応する近似図形内の輝度を増加させたが、この場合には、例えば、特徴点が 3 以上あるならば、それらを頂点とする多角形を生成して、その内部の輝度を増加させるようにしてもよい。また、特徴点が 2 以上あるならば、それらを代表点とする何らかの図形を発生させて、その内部の輝度を増加させるようにしてもよい。あるいは、各特徴点を中心とするある程度の大きさをもつ円などの図形を発生させ、発生させた各図形を、輝度、色、点滅などにより目立つようにして表示するようにしてもよい。

【0258】前述した図 35 では、物体に対応する近似図形の両軸方向への射影をバーの形で表示したが、この場合には、例えば、特徴点が 3 以上あるならば、それらを頂点とする多角形を生成して、該多角形の両軸方向への射影をバーの形で表示するようにしてもよい。また、特徴点が 2 つであるならば、それらを頂点とする矩形も

40

しくは長方形を生成して、その両軸方向への射影をバーの形で表示するようにしてもよい。一方、特徴点が 1 つであるならば、前述した図 36 のような方法で表示することができる。

【0259】前述した図 38 では、近似図形をもとにするなどして物体の画像を切り出してアイコンとして表示したが、この場合も、例えば、特徴点をもとにして物体の画像を切り出してアイコンとして表示することが可能である。

10 【0260】前述した図 39 では、近似図形をマップ表示したが、この場合も、例えば、前述のように特徴点をもとにして発生した何らかの図形をマップ表示することが可能である。

【0261】なお、図 37、図 40、図 41 の方法は、この場合も同様に可能である。

【0262】以上、いくつかのバリエーションを例示したが、これらの他にも、関連情報の表示等が可能な物体の領域を明示的に示す方法や、指示をしやすくする方法など、ユーザの利用を支援する方法は、様々なものが考えられる。

【0263】（第 5 の実施形態）次に、第 5 の実施形態として、映像中の物体を検索する方法について説明する。

【0264】本実施形態のブロック構成は例えば第 3 の実施形態（図 29）と同様である。なお、例えば図 29 において、関連情報表示部がない構成も可能であり（例えば、物体の検索自体を目的とし、関連情報を用いないシステムの場合）、また指示入力部がない構成も可能である（例えば、検索の指示に GUI を用いない場合）。もちろん、本実施形態もソフトウェアによっても実現可能である。

【0265】第 3 の実施形態では、楕円を用いる際に代表点を 2 つの焦点と楕円上の一点に取る場合について説明したが、ここでは、楕円の代表点として楕円の外接矩形の 3 頂点を用いる場合を例に取って説明する。もちろん、どのような代表点の取り方を用いても同じようにハイパーメディア検索が実現可能である。

【0266】なお、以下の $V_1 \sim V_4$ 、 P 、 Q 、 F_1 、 F_2 、 C_0 、 T 、 U 、 C はベクトル量である。

40 【0267】本発明では、物体の領域の軌跡を記述しているために、物体の通過した点、通過しなかった点を推定することにより物体の検索ができる。例えば、「この交差点の真ん中を通過してこちらの車線に走り去った車を検索せよ」とか、「ここから道路に進入したがこちらの車線には来なかった車を検索せよ」といった検索ができる。

【0268】図 42 や図 43 は、このような検索を行うための処理手順の例を示したものである。

50 【0269】図 42 は、物体の形の表現として矩形を用いた際の処理手順の一例である。

41

【0270】ここでは、通過した点または通過しなかった点としてQが指定されているものとする。

【0271】まず、ステップS100で、時間tに物体の出現した時間がセットされる。ステップS101では、ある時間tにおける矩形の代表点 V_1 、 V_2 、 V_3 の座標を取り出す。これは、対応する時間におけるスプライン関数の値として算出される。3つの矩形の頂点から残りの頂点の座標は、

$$V_4 = V_1 - V_2 + V_3$$

として簡単に求まる。

【0272】次に、ステップS102において、以下の式で表現される4つの関数の値を求める。

$$f_1(P) = (V_{2y} - V_{1y})(x - V_{1x}) - (V_{2x} - V_{1x})(y - V_{1y})$$

$$f_2(P) = (V_{2y} - V_{3y})(x - V_{2x}) - (V_{2x} - V_{3x})(y - V_{2y})$$

$$f_3(P) = (V_{3y} - V_{4y})(x - V_{3x}) - (V_{3x} - V_{4x})(y - V_{3y})$$

$$f_4(P) = (V_{1y} - V_{4y})(x - V_{4x}) - (V_{1x} - V_{4x})(y - V_{4y})$$

ただし、 $V_i = (V_{ix}, V_{iy})$ とする。

【0273】ステップS103では、求めた4つの関数に対して、

$$f_1(Q) f_3(Q) \leq 0 \quad \text{and} \quad f_2(Q) f_4(Q) \leq 0$$

が成立するかどうかを判定する。

【0274】成立する場合には、物体は時刻tに指定された点Qを通過することになるため、物体は点Qを通過*

$$a = |V_2 - V_1| / 2$$

$$b = |V_2 - V_3| / 2$$

$$F_1 = C_0 + e(V_2 - V_1) / 2 \quad (a > b \text{ のとき}),$$

$$C_0 + e(V_2 - V_3) / 2 \quad (a \leq b \text{ のとき})$$

$$F_2 = C_0 - e(V_2 - V_1) / 2 \quad (a > b \text{ のとき}),$$

$$C_0 - e(V_2 - V_3) / 2 \quad (a \leq b \text{ のとき})$$

ただし、 C_0 と e は次の通りである(e は a と b の大小 * 関係に応じて異なる)。

$$C_0 = (V_1 + V_3) / 2$$

$$e = \{\sqrt{(a^2 - b^2)}\} / a \quad (a > b \text{ のとき}),$$

$$\{\sqrt{(b^2 - a^2)}\} / b \quad (a \leq b \text{ のとき})$$

次に、ステップS113において、次の条件が満たされ *て異なる)。

るかどうかを判定する(条件は a と b の大小関係に応じ*

$$a > b \text{ のときの条件} \quad |F_1 - Q| + |F_2 - Q| \leq 2a$$

$$a \leq b \text{ のときの条件} \quad |F_1 - Q| + |F_2 - Q| \leq 2b$$

この条件が満たされている場合には、点Qは時刻tに楕円の内部にあることになる。従って、物体は点Qを通過したと判定し、処理を終了する(ステップS114)。一方、条件が満たされない場合には、点Qは時刻tにおいて楕円の外にあることになる。従って、他の時刻tについて同様の処理を続ける。

【0282】ステップS115は終了条件で、tが物体が画面から消失した時刻であるかどうかを判定する。消失時刻である場合には、物体は点Qを通過しなかったと

42

*すると判定する(ステップS104)。成立しない場合には、時刻tには点Qを通過しないため、他の時間において通過したかどうかを引き続き調べる。

【0275】ステップS105では、全ての時間tについて調査したかどうかを判定するため、tが物体が画面から消えた時間と同じかどうかを調べる。もしも同じならば処理を終了し、物体は点を通過しなかったと判定する(ステップS107)。tが物体が消えた時間よりも小さければ、ステップS106においてtを1だけ増加させてS101からの処理を繰り返す。

【0276】以上の処理を検索対象となっている全ての物体に対して行うことにより、検索条件を満たす物体を検索できる。

【0277】図43は、物体の形の表現として楕円を用いた際の処理手順の一例である。

【0278】まず、ステップS110で、時間tに物体の出現した時間がセットされる。

【0279】次に、ステップS111で、ある時間tにおける楕円の代表点 V_1 、 V_2 、 V_3 の座標を取り出す。これら代表点は、楕円の外接矩形の3頂点であり、1つの辺を挟んで時計回りに V_1 、 V_2 、 V_3 と連続して並んでいる。これを算出する処理は矩形に対して行うと同様の処理である。

【0280】次に、ステップS112で、以下の式で表される a 、 b 、および点 F_1 、 F_2 を求める(F_1 、 F_2 は a と b の大小関係に応じて以下のように求める)。

【0281】

【0281】

$$(a > b \text{ のとき}),$$

$$(a \leq b \text{ のとき})$$

$$(a > b \text{ のとき}),$$

$$(a \leq b \text{ のとき})$$

*関係に応じて異なる)。

$$C_0 = (V_1 + V_3) / 2$$

$$e = \{\sqrt{(a^2 - b^2)}\} / a \quad (a > b \text{ のとき}),$$

$$\{\sqrt{(b^2 - a^2)}\} / b \quad (a \leq b \text{ のとき})$$

*て異なる)。

判定し、処理を終了する(ステップS117)。tが消失時刻でなければ、ステップS116でtを増加させ、ステップS111から処理を繰り返す。

【0283】ここまでの処理を全ての検索対象の物体に対して行うことにより、検索条件を満たす物体を検索できる。

【0284】なお、上記では、指定点が近似図形に含まれるか否かを通過したか否かの判断基準としたが、例えば、指定点が近似図形の近傍に位置した場合も通過した

とみなす場合や、指定点が一定フレーム以上連続して近似図形に含まれたときに始めて通過したと判断する場合など、種々の基準が考えられる。

【0285】なお、物体の形の表現として他の図形が用いられる場合も図形に応じた同様の処理を行うことによって、検索条件を満たす物体を検索することができる。

【0286】また、複数の通過点あるいは複数の非通過点が指定されている場合には、それら全ての指定点に対して同じ処理を行えばよい。もちろん、1または複数の通過点と1または複数の非通過点を混在して指定することも可能である。また、複数の通過点または非通過点について組み合わせ論理を用いて検索することも可能である。例えば、「a点またはb点のいずれか一方のみを通過し、かつ、c点およびd点のいずれも通過していない物体を検索せよ」といった検索も可能である。

【0287】ところで、通過点による検索は、物体が通過点にどのくらい存在していたかで検索ができるように拡張することができる。このような検索には、「10分以上立ち読みしていた人を検索せよ」とか「キャッシュディスクの前に3分以上いた人を検索せよ」などがある。このような検索を行うには、指定された位置に物体が存在していた時間をカウントし、ユーザにより指定された時間より多くの時間となった物体だけを検索結果として提示すればよい。

【0288】さらに別の検索拡張例として、物体の大きさ（面積）による条件を付加した場合について説明する。

【0289】物体の形状が矩形もしくは楕円で表現されている場合、ある時刻tにおける物体の面積は、矩形の場合には、

$$S_R = |V_2 - V_1| \cdot |V_3 - V_2|$$

楕円の場合には、

$$S_E = a b \pi$$

により計算できる。

【0290】この値を用いて、物体が指定点を通過するか否かだけでなく、面積が S_S 以上 S_L 以下というような条件を用いて検索ができる。例えば、「道を歩く人を検索せよ。ただし、犬や猫は検索しないこと」といった要求に対して、通常の犬や猫よりも大きな面積を指定しておくことにより検索の精度を向上させることができる。

【0291】次に他の検索例として、似たような軌跡を描いて移動した物体の検索の方法について説明する。

【0292】第1の物体の軌跡と第2の物体の軌跡をそれぞれT、Uとする。第1の物体の存在時間と第2の物体の存在時間をそれぞれ N_T 、 N_U とする。ただし、 $N_T \geq N_U$ が成り立っているものとする。さらに、どちらの物体の出現時間も $t=0$ であるものとする。これらの条件は、適当にTとUを入れ替えて、時間軸の原点をずらすことにより常に満たすことができる。

【0293】このとき、TとUの距離 $d(T, U)$ を以下

のように定義する。

【0294】

【数1】

$$d(T, U) = \min_{0 \leq i \leq N_T - N_U} \sum_{j=0}^{N_U} E^2(T(j), U(j+i))$$

【0295】ここで、tにおけるTの座標を $T(t)$ と表し、 $E(P, Q)$ はユークリッド距離を表すものとした。

【0296】この軌跡間の距離を用いて、ユーザによって指定された物体の軌跡と他の物体の軌跡との距離を、他の全ての物体についてそれぞれ算出し、最も距離の短い物体を提示するか、または距離の短い物体をユーザにより指定された数だけで提示することで、似ている軌跡を描く物体の検索が行える。

【0297】また、ユーザによりマウス等の入力装置により描かれた軌跡と類似した軌跡を描く物体を検索することも可能である。この場合には、ユーザにより描かれた軌跡には時間情報が含まれていないため、 $d(T, U)$ とは別の方法で軌跡間の距離を算出する必要がある。そのためには、軌跡Tとユーザにより描かれた軌跡Uとの距離 $d'(T, U)$ を次のように計算する。

【0298】

【数2】

$$d'(T, U) = \sum_{i=0}^{N_{PU}-1} \min_{0 \leq j \leq N_T} E^2(T(j), U_i)$$

【0299】ここで、ユーザにより描かれた軌跡を点列 U_i ($0 \leq i < N_{PU}$)により表す。 N_{PU} は点列の数である。この距離の短い物体を、類似した軌跡を描く物体として1つまたは数個提示することで検索が行える。

【0300】また、物体の中心の軌跡が記述されている場合には、これらの軌跡をTおよびUとして $d(T, U)$ の短い物体を検索すればよい。ただし、物体の形を近似した矩形や楕円の軌跡のみの情報しか得られない場合には、中心の軌跡を推定してから物体の軌跡間の距離を算出する。ある時刻tにおける中心Cの推定値は、矩形または楕円の頂点の座標 V_1, V_2, V_3 から次のように推定する。

【0301】

【数3】

$$C \approx \frac{V_1 + V_3}{2}$$

【0302】この推定により、全ての物体の軌跡の中から、類似した軌跡を検索することができる。

【0303】なお、上記では、物体の領域に対する近似図形の代表点を用いる場合を例として説明したが、第2の実施形態のように物体の領域の特徴点を用いる場合にも同様に可能である。この場合に、物体が指定点を通過したか否かの判断は、例えば、特徴点と指定点との距離が基準値以下か否かによって行えばよい。

【0304】なお、以上の各実施形態あるいは各構成は、任意に組み合わせて実施可能である。

【0305】以上説明した各構成は、コンピュータに所定の手段を実行させるための（あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるための）プログラムを記録したコンピュータ読取り可能な記録媒体としても実施することもできる。

【0306】本発明は、上述した実施の形態に限定されるものではなく、その技術的範囲において種々変形して実施することができる。

【0307】

【発明の効果】本発明によれば、映像中における対象となる物体の領域を該物体の領域に対する近似図形の代表点または該物体の領域の特徴点の位置データをフレームの進行に沿って並べたときの軌跡を近似した関数のパラメータとして記述することにより、映像中の所望の物体の領域を少ないデータ量で記述でき且つその作成やそのデータの扱いも容易にすることができる。

【0308】また、本発明によれば、ユーザによる映像中の物体の指示やその判定を容易にすることができる。

【0309】また、本発明によれば、映像中の物体の検索を容易にすることができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係る物体領域データ変換装置の構成例を示す図

【図2】映像中の物体の領域を物体領域データで記述するための処理の概要を説明するための図

【図3】物体の領域を楕円で近似する例について説明するための図

【図4】物体の領域を近似する楕円の代表点を求める例について説明するための図

【図5】物体領域データのデータ構造の一例を示す図

【図6】物体領域データ中の近似図形データのデータ構造の一例を示す図

【図7】近似図形データ中の代表点軌跡データのデータ構造の一例を示す図

【図8】近似図形を平行四辺形とした場合の代表点の例について説明するための図

【図9】近似図形を多角形としたときの代表点の例について説明するための図

【図10】同実施形態における処理手順の一例を示すフローチャート

【図11】映像中の物体の領域を複数の楕円で表現した例を示す図

【図12】複数の近似図形データを含む物体領域データのデータ構造の一例を示す図

【図13】映像中の物体の領域を物体領域データで記述するための他の処理の概要を説明するための図

【図14】近似矩形を求める処理手順の一例を示すフ

ローチャート

【図15】傾斜した細長い物体を傾斜を持たない矩形で近似した様子を示す図

【図16】物体をその傾斜に応じた傾斜を持つ矩形で近似した様子を示す図

【図17】近似矩形を求める処理手順の他の例を示すフローチャート

【図18】近似矩形から近似楕円を求める方法について説明するための図

【図19】近似矩形から近似楕円を求める処理手順の一例を示すフローチャート

【図20】隣接するフレーム間で近似図形の代表点同士を対応付ける方法について説明するための図

【図21】隣接するフレーム間で近似図形の代表点同士を対応付けるための処理手順の一例を示すフローチャート

【図22】物体領域データのデータ構造の他の例を示す図

【図23】図形種IDと図形の種類と代表点数との対応の一例を示す図

【図24】関数IDと関数形式と関数パラメータと制約条件との対応の一例を示す図

【図25】関連情報のデータ構造の具体例を示す図

【図26】関連情報のデータ構造例の他の具体例を示す図

【図27】本発明の第2の実施形態に係る物体領域データ変換装置の構成例を示す図

【図28】同実施形態における処理手順の一例を示すフローチャート

【図29】本発明の第3の実施形態に係る情報処理装置の構成例を示す図

【図30】同実施形態における処理手順の一例を示すフローチャート

【図31】物体領域データを利用したハイパーメディア・コンテンツの表示例を示す図

【図32】同実施形態における処理手順の他の例を示すフローチャート

【図33】本発明の第4の実施形態における処理手順の一例を示すフローチャート

【図34】関連情報を持つ物体領域の表示を変更する例を示す図

【図35】関連情報を持つ物体領域の位置を表示する一例を示す図

【図36】関連情報を持つ物体領域の位置を表示する他の例を示す図

【図37】関連情報を持つ物体領域の説明リストを表示する例を示す図

【図38】関連情報を持つ物体領域をアイコン表示する例を示す図

【図39】関連情報を持つ物体領域の位置をマップ表示

47

する例を示す図

【図 4 0】 物体領域の指定を容易にするために画像再生速度を制御する例を示す図

【図 4 1】 物体領域の指定を容易にするために画像キャプチャを可能とする例を示す図

【図 4 2】 本発明の第 5 の実施形態における処理手順の一例を示すフローチャート

【図 4 3】 同実施形態における処理手順の他の例を示すフローチャート

【符号の説明】

1 0 0, 2 3 0 … 映像データ記憶部

1 0 1 … 領域抽出部

48

* 1 0 2 … 領域の図形近似部

1 0 3 … 図形代表点抽出部

1 0 4 … 代表点の曲線近似部

1 0 5, 2 3 5 … 関連情報記憶部

1 0 6, 2 3 6 … 領域データ記憶部

2 3 3 … 特徴点抽出部

2 3 4 … 特徴点の曲線近似部

3 0 1 … 映像データ表示部

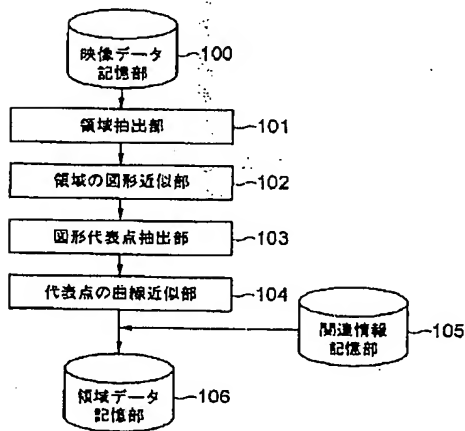
3 0 2 … 制御部

10 3 0 3 … 関連情報表示

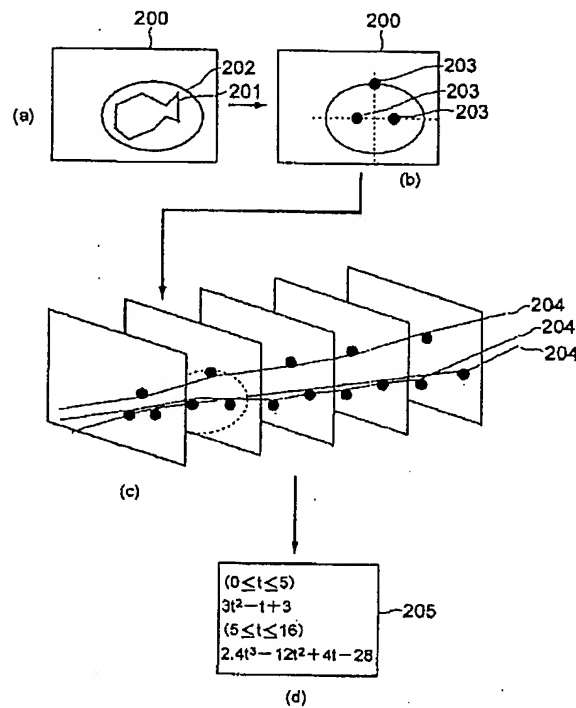
3 0 4 … 指示入力部

*

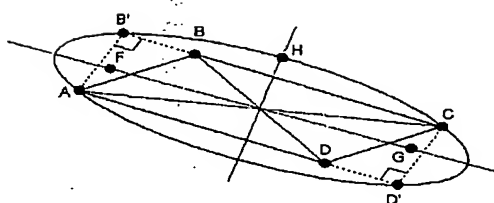
【図 1】



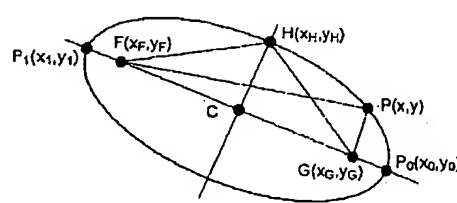
【図 2】



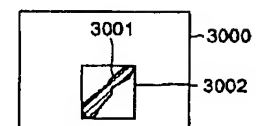
【図 3】



【図 4】



【図 15】



【図 5】

物体領域データ	
ID番号	400
先頭フレーム番号	401
最終フレーム番号	402
関連情報へのポイント	403
近似図形数	404
近似図形データ	405

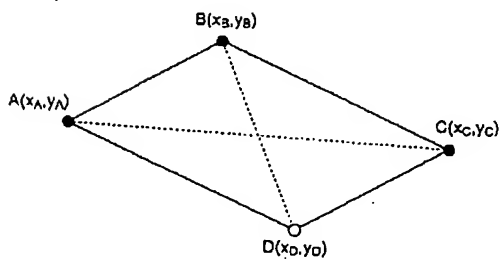
【図 6】

近似図形データ	
図形種ID	1300
代表点数	1301
代表点軌跡データ(1X)	1302
代表点軌跡データ(1Y)	1303
代表点軌跡データ(2X)	1403
代表点軌跡データ(2Y)	
⋮	
代表点軌跡データ(MX)	
代表点軌跡データ(MY)	

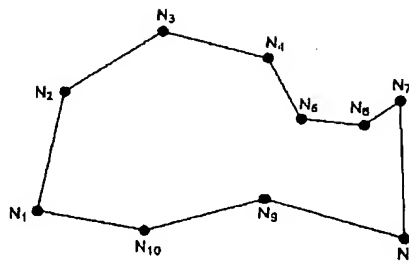
【図 7】

代表点軌跡データ	
節点フレーム番号	1400
多項式次数	1401
多項式係数(0次)	1402
多項式係数(1次)	
⋮	
多項式係数(K次)	
節点フレーム番号	
多項式次数	
多項式係数(0次)	
多項式係数(1次)	
⋮	
多項式係数(K'次)	

【図 8】



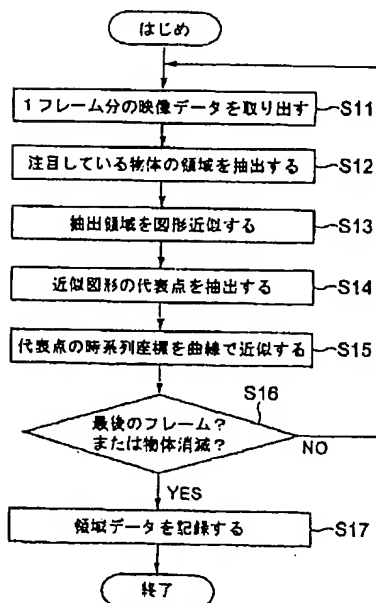
【図 9】



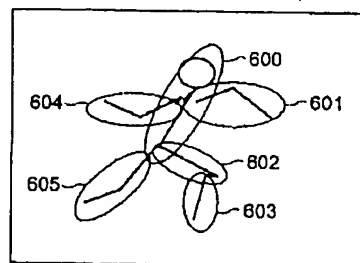
【図 12】

物体領域データ	
ID番号	400
先頭フレーム番号	401
最終フレーム番号	402
関連情報へのポイント	403
近似図形数	404
近似図形データ1	405
近似図形データ2	
⋮	
近似図形データL	

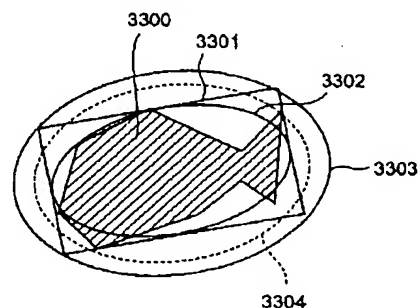
【図 10】



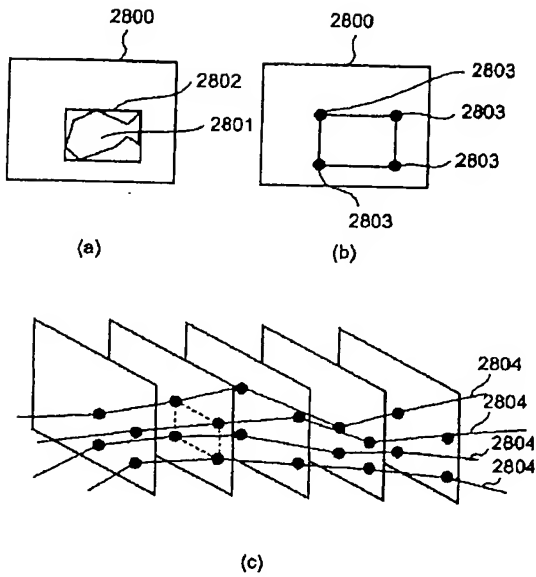
【図 11】



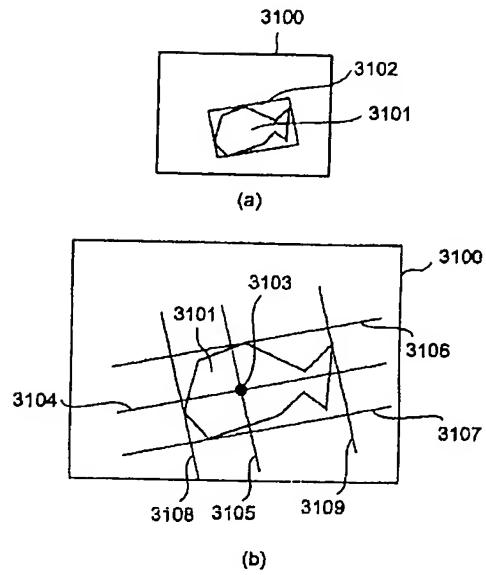
【図 18】



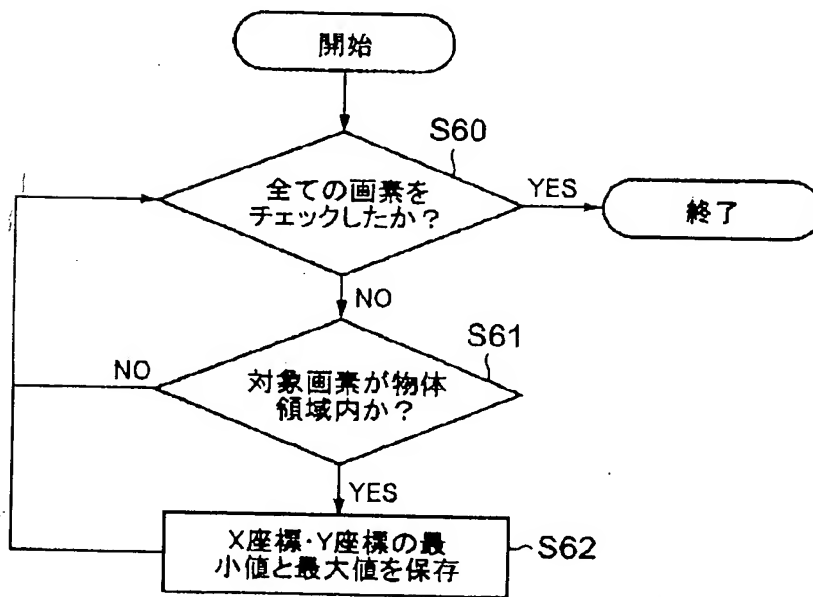
【図13】



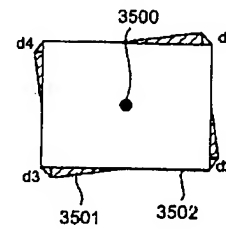
【図16】



【図14】



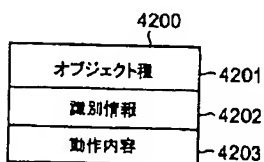
【図20】



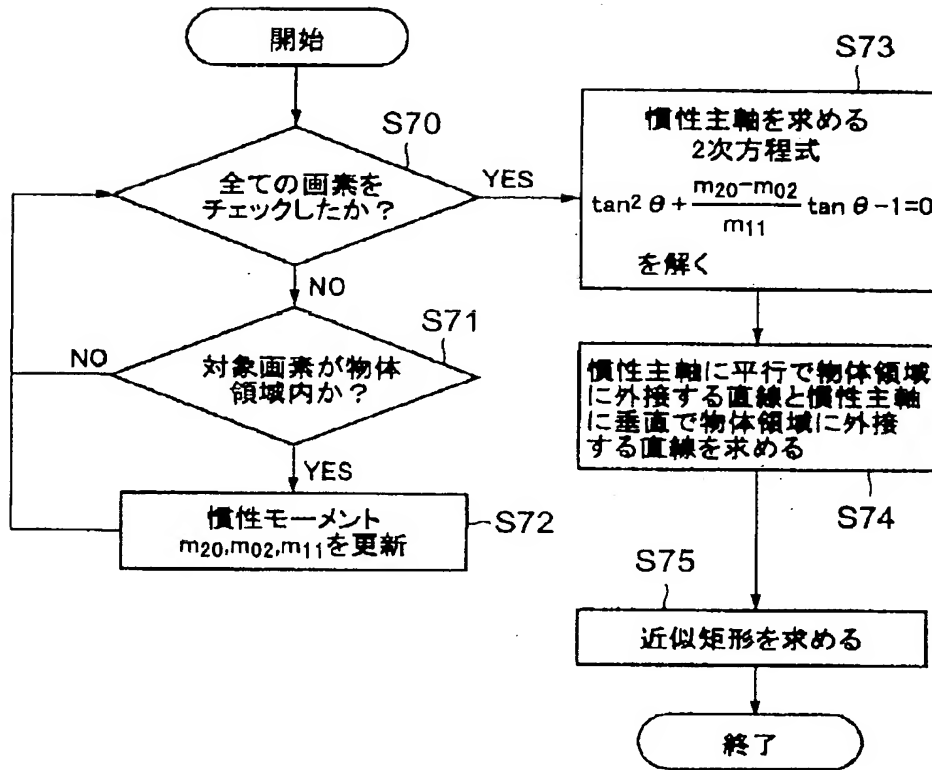
【図23】

図形種ID	図形の種類	代表点数
0	FORBIDDEN	-
1	CENTROID	1
2	RECTANGLE	3
3	RECTANGLE & CENTROID	4
4	ELLIPSE	3
5	ELLIPSE & CENTROID	4

【図25】



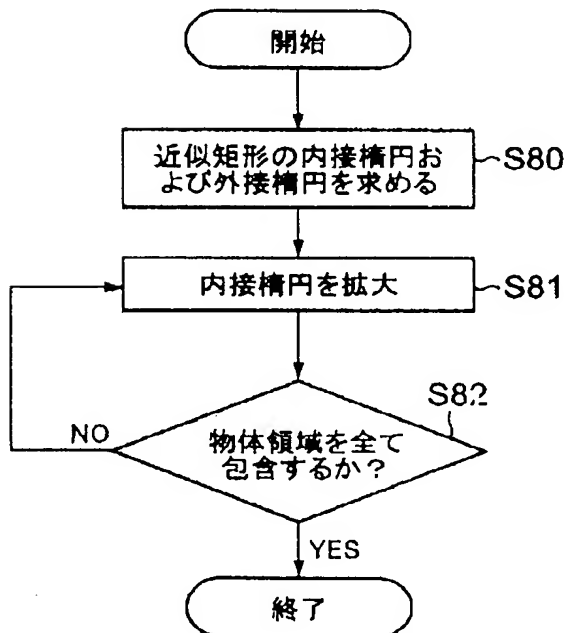
【図17】



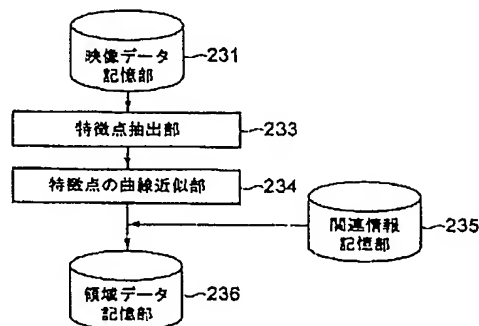
【図26】

4300	
名前	4301
著作権情報	4302
コピー許可情報	4303
フットマーク	4304
関連情報URL	4305
アクセス制限情報	4306
課金情報	4307
注釈	4308

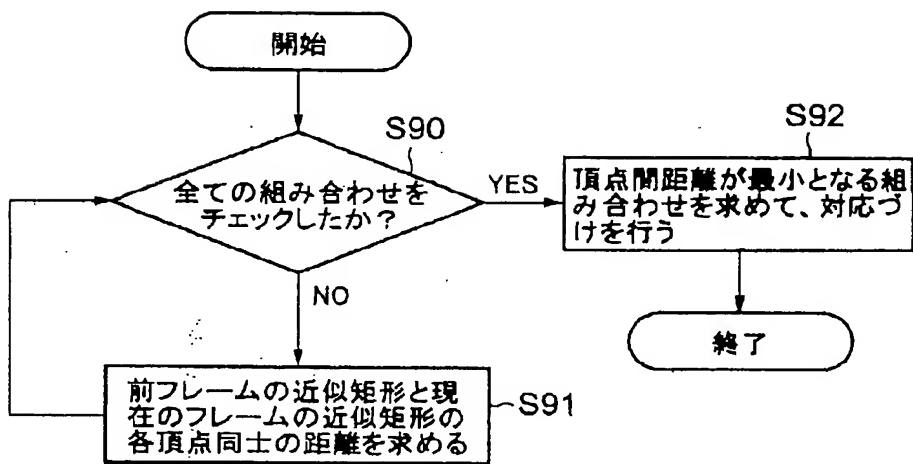
【図19】



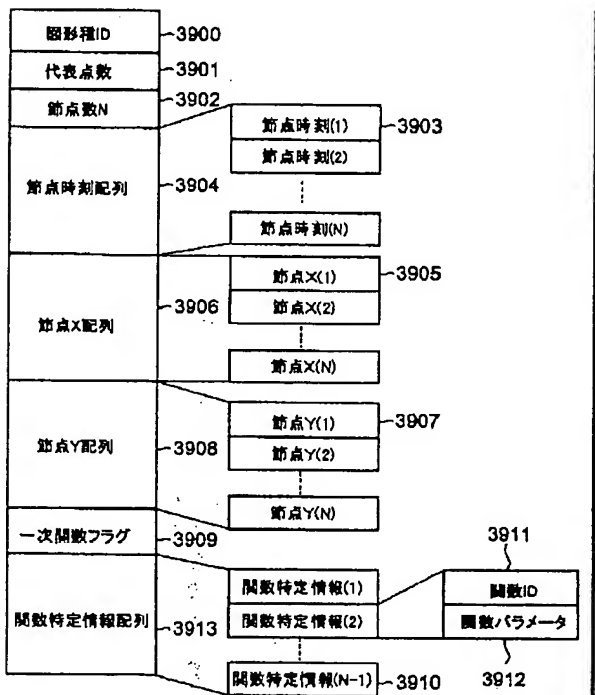
【図27】



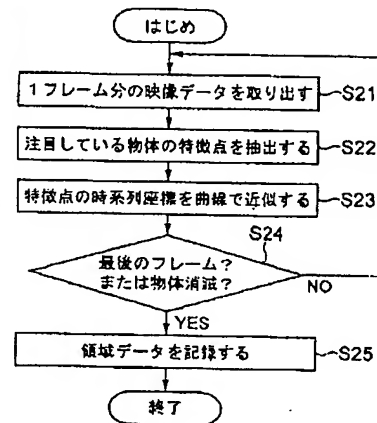
【図21】



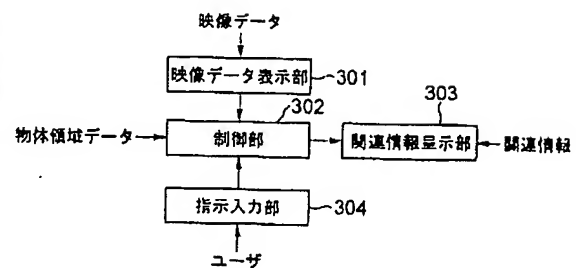
【図22】



【図28】



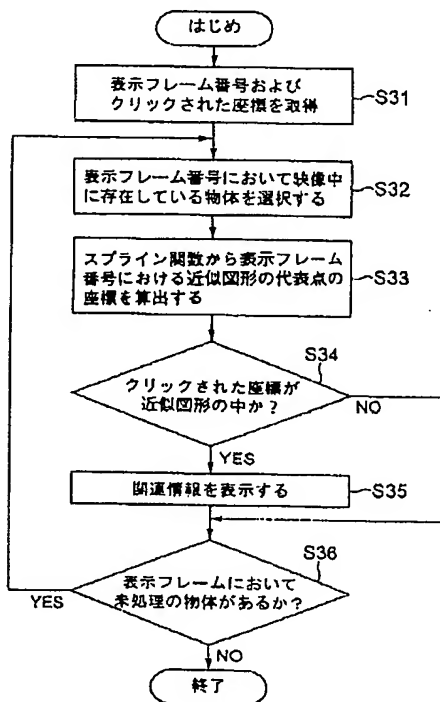
【図29】



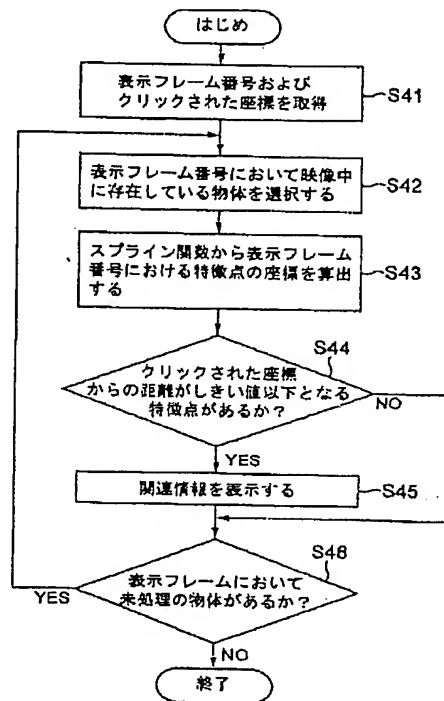
【図 24】

関数ID	関数形式	関数 パラメータ	制約条件
-1	なし(関数は定義されない)	なし	なし
0	$f(t)=f_a+v_a(t-t_a)$	なし	$v_a=\frac{f_b-f_a}{t_b-t_a}$
1	$f(t)=f_a+v_a(t-t_a)+\frac{1}{2}a_a(t-t_a)^2$	a_a	$v_a=\frac{f_b-f_a}{t_b-t_a}-\frac{1}{2}a_a(t_b-t_a)$

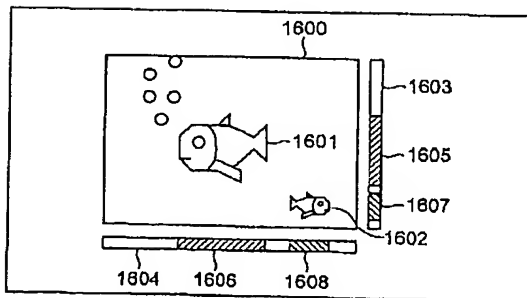
【図 30】



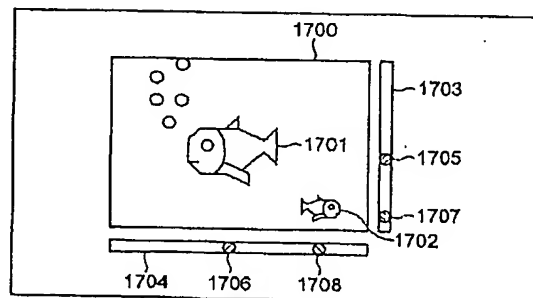
【図 32】



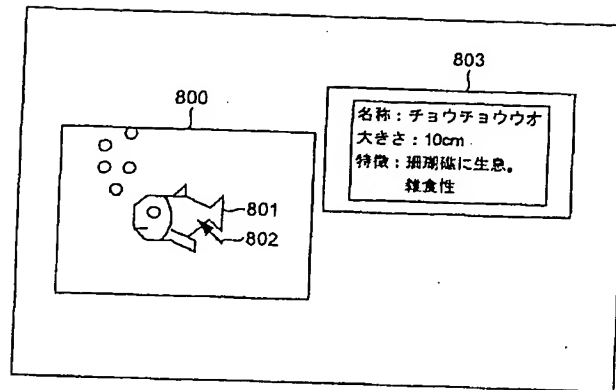
【図 35】



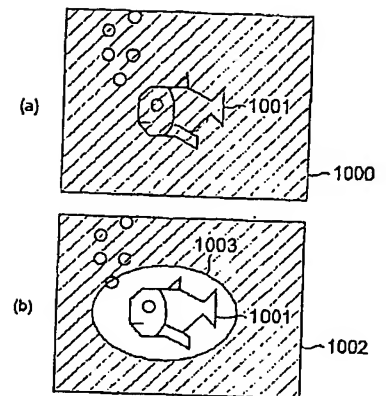
【図 36】



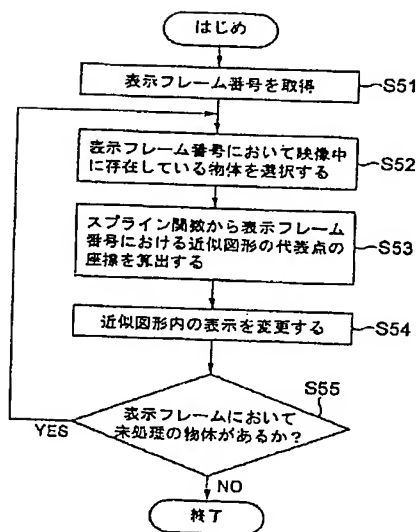
【図31】



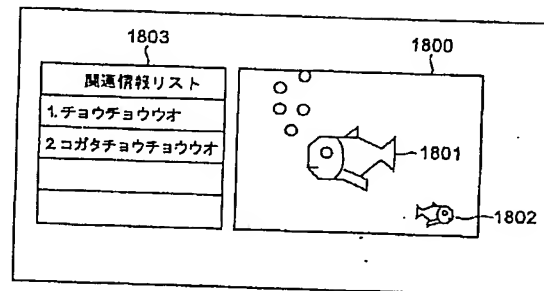
【図34】



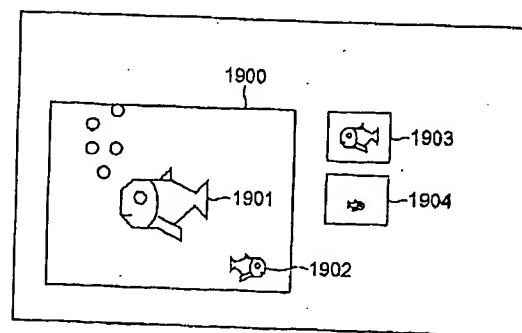
【図33】



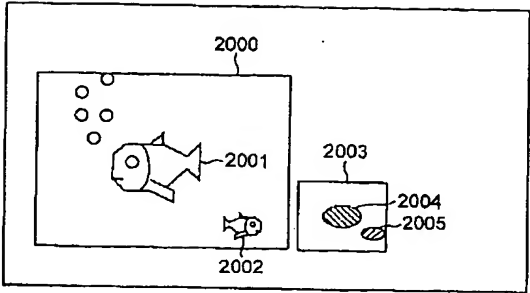
【図37】



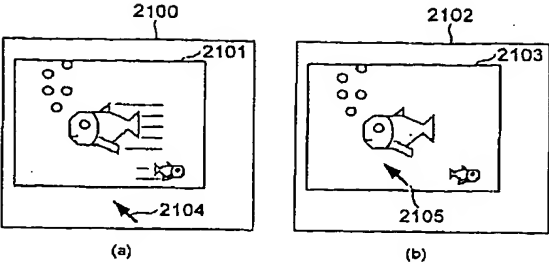
【図38】



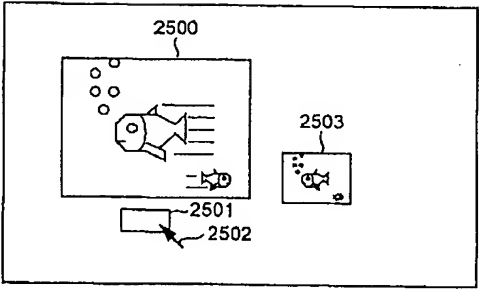
【図 39】



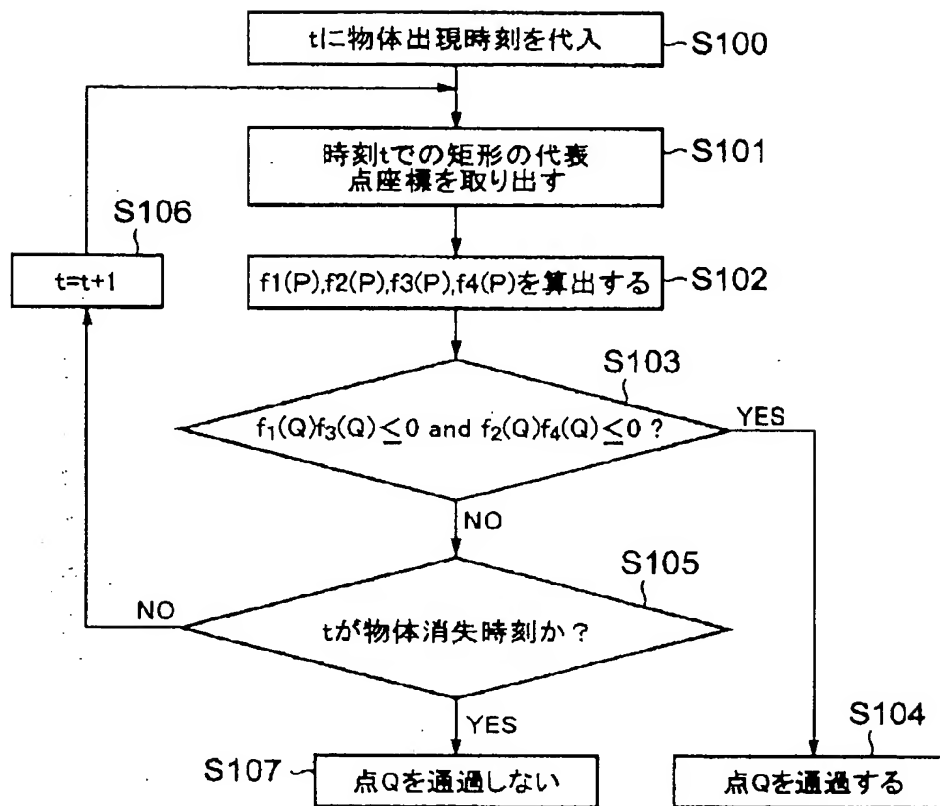
【図 40】



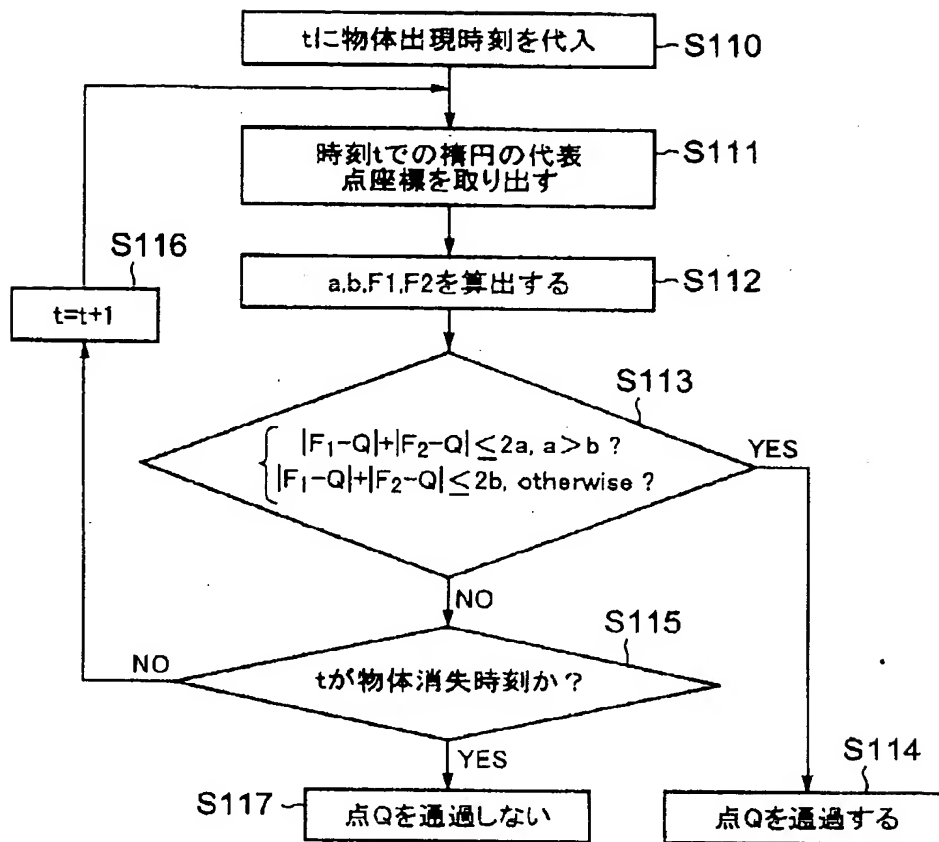
【図 41】



【図42】



【図 4 3】



フロントページの続き

(72) 発明者 三田 雄志
 神奈川県川崎市幸区小向東芝町 1 番地 株
 式会社東芝研究開発センター内
 (72) 発明者 山本 晃司
 神奈川県川崎市幸区小向東芝町 1 番地 株
 式会社東芝研究開発センター内

F ターム (参考) 5B050 BA07 BA13 BA18 DA08 FA02
 FA09 GA08
 5B075 ND12 NK39 PP03 PQ02
 5C023 AA06 AA11 AA18 BA11 DA01
 5C059 KK37 MB03 MB22 SS11